

TRAITÉ

DU

CALCUL DES PROBABILITÉS.



PARIS. — IMPRIMERIE DE GAUTHIER-VILLARS,
(832) Quai des Augustins, 55.

TRAITÉ

DU

CALCUL DES PROBABILITÉS,

PAR H. LAURENT,

RÉPÉTITEUR D'ANALYSE A L'ÉCOLE POLYTECHNIQUE, MEMBRE DU CERCLE DES ACTUAIRES FRANÇAIS.





PARIS,

GAUTHIER-VILLARS, IMPRIMEUR-LIBRAIRE

DU BUREAU DES LONGITUDES, DE L'ÉCOLE POLYTECHNIQUE,

SUCCESSEUR DE MALLET-BACHELIER,

Quai des Augustins, 55.

1873



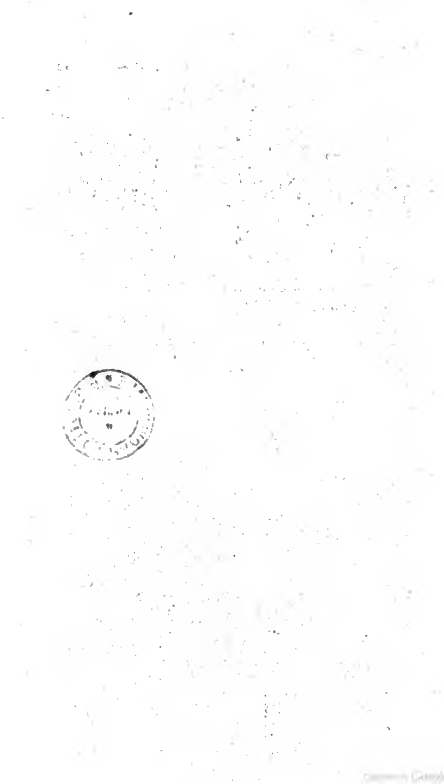


TABLE DES MATIÈRES.

CHAPITRE IV.

SUR LES MÉTHODES DANS LES SCIENCES D'OBSERVATION.

CHAPITRE V.

SUR LES COMPAGNIES D'ASSURANCES.

NOTES.

ERRATA.

Page 51, ligne 14, *au lieu de :* égal à, *lisez :* est complémentaire de.

Page 52, ligne 8, *ajoutez :* ces probabilités seraient du reste les mêmes si l'on admettait le zéro comme nombre pair.

Page 85. La dernière ligne commence l'énoncé d'un problème qu'il faut compléter comme il suit :

On suppose que toutes les combinaisons qui peuvent se présenter en deux coups de filet soient également possibles.

Page 103, ligne 17, *au lieu de :* d'où, *lisez :* où.

Page 186, ligne 12, *au lieu de :* résumés, *lisez :* mesurés.

Page 210, ligne 12, *au lieu de :* espérée à la mort, *lisez :* espérée en cas de vie.



PRÉFACE.

Les personnes qui désirent étudier le Calcul des Probabilités éprouvent généralement des difficultés qui tiennent moins à la nature du sujet qu'à l'absence de Traités réellement classiques. Et, en effet, pour aborder la célèbre *Théorie analytique des probabilités* de Laplace, il faut déjà être, jusqu'à un certain point, familiarisé avec l'analyse des hasards, l'auteur ne traitant, comme il l'avoue lui-même, que les questions les plus difficiles; quant aux Livres très-estimés de MM. Lacroix et Cournot, ils sont trop élémentaires, même pour servir d'Introduction à la lecture du Traité de Laplace. Il n'en est pas moins vrai que c'est par ces deux derniers Ouvrages qu'il convient d'aborder l'étude des Probabilités. Le Traité de Poisson, sur la probabilité des jugements, écrit, comme son titre l'indique, dans un but spécial, contient cependant une Introduction qui peut être considérée aujourd'hui comme le meilleur Traité réellement classique; malheureusement, il est fort incomplet.

Il m'a semblé que je rendrais service à un bon nombre de personnes, et principalement aux officiers d'artillerie et aux candidats actuaires, en publiant un Traité très-élémentaire dans ses principes, mais cependant assez complet pour permettre de lire tout ce qui a été écrit sur la matière. Le présent Traité peut être considéré comme une véritable

Introduction au Traité de Laplace, quoiqu'il forme à lui seul un corps assez étendu pour embrasser toute la science des hasards. On regrettera peut-être que je n'aie point parlé de l'espérance morale, ni de la probabilité des jugements ou des témoignages; je dois avouer que ces théories, malgré l'autorité des savants qui les ont propagées, ne m'ont pas semblé présenter un caractère assez scientifique; et d'ailleurs il n'est heureusement pas besoin d'avoir recours à la considération métaphysique de l'espérance morale pour expliquer les paradoxes analogues à ceux que présente le jeu de Saint-Pétersbourg. Quant à la *probabilité d'un témoignage*, je n'ai jamais compris ce que pouvait signifier une pareille locution, le mot *probabilité* impliquant en soi l'idée de cas possibles et également possibles auxquels un événement relatif à un témoignage ne peut évidemment pas se rapporter.

Le Chapitre Ier de ce Traité contient quelques théories analytiques dont on fait un fréquent usage dans le Calcul des Probabilités; les personnes qui ne voudront pas pousser bien loin l'étude de la théorie des hasards pourront se contenter de lire le Chapitre II, qui est consacré au développement des principes fondamentaux et de leurs applications immédiates; le Chapitre III contient les théories qui dépendent de l'évaluation des grands nombres. Je crois avoir présenté ces théories avec plus de rigueur qu'on ne l'a fait jusqu'ici, en ce sens que j'ai toujours évalué une limite précise de l'erreur commise en bornant à leurs premiers termes les séries qui servent au calcul approché des grands nombres.

Le Chapitre IV contient la théorie des erreurs d'observa-

tion et du tir à la cible, l'étude des phénomènes et de leurs causes. J'ai cru devoir exposer la méthode des moindres carrés, en suivant la marche indiquée par Cauchy, tout en la modifiant, d'après les indications que M. Bienaymé a publiées dans son Mémoire inséré au tome XXIII du *Journal de Liouville*, Mémoire célèbre dont il m'a été impossible de donner une analyse à cause de sa grande étendue. On me reprochera peut-être de n'avoir pas fait d'applications numériques de la méthode des moindres carrés. J'ai regardé cette tâche comme impossible, et voici pour quels motifs : la méthode des moindres carrés n'a de valeur sérieuse qu'autant que l'on a à sa disposition un nombre immense d'observations ; or ces observations pouvaient être fictives ou réelles ; mais il eût été vraiment oiseux de faire des calculs très-longs et très-pénibles sur des nombres fictifs présentant certainement des circonstances en désaccord avec ce qui a lieu dans l'ordre naturel des choses. Traiter un exemple d'après les observations de quelque physicien en renom me paraissait également inutile. En effet, dans ses applications, la méthode des moindres carrés présente une variété infinie, et les personnes qui voudront s'en servir feront bien de ne pas s'en tenir à la lecture d'un exemple isolé, le cas qu'ils auront à traiter pouvant différer considérablement de ceux qui sont consignés dans les livres. On voit donc que, si j'avais voulu traiter des exemples numériques, j'aurais été forcé de donner trop d'importance à un Chapitre déjà très-chargé.

Le dernier Chapitre est consacré aux opérations des Compagnies d'assurances. Je n'ai rien à en dire, si ce n'est qu'il est rédigé à un point de vue tout à fait pratique ; les for-

mules que j'indique sont celles dont on fait réellement usage dans le calcul des tarifs que l'on distribue au public. Je dois dire cependant qu'il m'a été impossible d'admettre les anciennes formules relatives aux assurances de survie, que j'ai toujours considérées comme fausses.

Notre Traité se termine par une liste des principaux Ouvrages relatifs au Calcul des probabilités; je dois, malgré les renseignements que j'ai puisés de tous côtés, avoir omis bien des noms. Je demande mille fois pardon aux auteurs que je n'ai pas cités : tout ce que je puis faire, c'est de les prier de vouloir bien m'adresser leurs réclamations, auxquelles je donnerai satisfaction dans une autre édition.

Un des Membres les plus distingués du Cercle des Actuaires français, M. Marc Achard, ayant bien voulu se charger de la révision des épreuves de cet Ouvrage, je lui adresse ici mes remercîments (*).

(*) Les géomètres disent une droite, une développable, etc., pour une ligne droite, une surface développable, etc. J'ai cru pouvoir imiter ce langage incorrect et éviter la répétition du mot *boule* en écrivant une rouge, une blanche, au lieu d'une boule rouge, une boule blanche, etc.

TRAITÉ

DU

CALCUL DES PROBABILITÉS.

CHAPITRE PREMIER.

ANALYSE COMBINATOIRE.

Arrangements, permutations, combinaisons.

On appelle *arrangements* de m objets n à n les résultats obtenus en prenant n de ces objets de toutes les manières possibles, de telle sorte que ces résultats diffèrent soit par l'ordre, soit par la nature de ces objets. On désigne habituellement par A_m^n le nombre des arrangements de m objets pris n à n.

Supposons que l'on connaisse A_m^n, cherchons à former A_m^{n+1}; pour former les arrangements $n+1$ à $n+1$ de m objets, il suffit de placer *à la suite* de chaque arrangement n à n chacun des objets qui n'y entrent pas, ce qui fournit $(m-n)A_m^n$ résultats tous différents soit par la nature du dernier objet, soit par les arrangements n à n qui ont servi à les former. Il est facile de voir, du reste, qu'un arrangement quelconque $n+1$ à $n+1$ se trouve compris dans les résultats de notre opération, car on peut le former en plaçant un dernier objet à la suite des n autres, qui forment

un arrangement n à n. On a donc

$$A_m^{n+1} = A_m^n(m-n),$$

d'où l'on conclut aisément

$$A_m^n = m(m-1)\ldots(m-n+1),$$

en observant que $A_m^1 = m$, que, par suite, $A_m^2 = m(m-1)\ldots$.

On appelle *permutations* de m objets les résultats obtenus en groupant ces m objets de toutes les manières possibles, de telle sorte que tous les résultats diffèrent par l'ordre des objets qu'ils contiennent. Il résulte de là que, si l'on désigne par P_m le nombre des permutations de m objets, on a

$$P_m = A_m^m = m(m-1)\ldots3.2.1 = m!,$$

en adoptant la notation très-répandue $m!$ pour désigner le produit $1.2.3\ldots m$.

On appelle *combinaisons* de m objets pris n à n les résultats obtenus en groupant n de ces m objets de toutes les manières possibles, de telle sorte que tous les résultats diffèrent seulement par la nature des objets qu'ils contiennent; en d'autres termes, deux résultats contenant à l'ordre près les mêmes objets sont censés identiques. Si l'on désigne par C_m^n le nombre de combinaisons de m objets pris n à n, il est facile de voir que l'on a

$$C_m^n = \frac{A_m^n}{P_n} = \frac{m(m-1)\ldots(m-n+1)}{1.2.3\ldots n} = \frac{m!}{(m-n)!\,n!} = C_m^{m-n},$$

et que si, par suite, on fait $\alpha + \beta = m$, on a

$$C_m^\alpha = C_m^\beta = \frac{m!}{\alpha!\,\beta!}.$$

En effet, supposons que l'on connaisse C_m^n; pour obtenir A_m^n il suffira, dans chaque combinaison, de permuter les n objets qu'elle renferme, car les combinaisons sont des arrangements considérés comme identiques, lorsqu'ils ne diffèrent que par l'ordre des objets; on a donc

$$C_m^n \times P_n = A_m^n,$$

d'où l'on déduit les formules précédentes.

Sur une méthode générale suivie dans les recherches d'Analyse combinatoire.

Le but de l'Analyse combinatoire est de calculer le nombre des groupes que l'on peut former en arrangeant des objets d'une manière déterminée. Ce nombre peut s'obtenir directement dans un grand nombre de cas, mais la plupart des auteurs ont substitué aux méthodes directes une marche qui conduit plus sûrement au résultat et à laquelle on est même forcé d'avoir recours dans un grand nombre de cas, où la recherche directe ne conduirait qu'à des résultats impossibles à calculer. C'est ainsi que le produit $1.2.3\ldots n$, qui représente le nombre des permutations de n objets, serait impossible à évaluer si n représentait un nombre tel que 1000 ou 10000, la longueur des opérations à faire devenant un obstacle presque insurmontable.

La méthode indirecte dont je veux parler consiste à considérer la quantité que l'on veut évaluer comme le terme général du développement d'une fonction en série ordonnée, suivant les puissances ascendantes d'une variable x. Pour donner tout de suite une idée de cette méthode, je vais en faire l'application à la recherche du nombre de combinaisons de m objets pris n à n.

Si l'on considère le produit $(1+ax)(1+bx)\ldots(1+lx)$, dans lequel le nombre des facteurs est m, il est clair que le coefficient de x^n sera la somme des produits obtenus en multipliant n des lettres $a, b, c, \ldots, l$ entre elles de toutes les manières possibles; ces produits sont en nombre égal à celui des combinaisons de m objets, pris n à n. Si donc on fait $a=b=c=\ldots=1$, le coefficient de x^n, dans $(1+x)^m$, sera précisément C_m^n.

Posant alors

$$(1) \qquad (1+x)^m = 1 + C_m^1 x + \ldots + C_m^n x^n + \ldots,$$

on a, en différentiant, par rapport à x,

$$m(1+x)^{m-1} = C_m^1 + \ldots + n C_m^n x^{n-1},$$

et, par suite,

$$(2) \qquad m(1+x)^m = (1+x)(C_m^1 + \ldots + n C_m^n x^{n-1});$$

si l'on multiplie ensuite l'équation (1) par m, et si l'on identifie le second membre avec celui de la formule (2), on trouve

$$m C_m^n = n C_m^n + (n+1) C_m^{n+1},$$

d'où l'on tire

$$C_m^{n+1} = \frac{m-n}{n+1} C_m^n.$$

Comme on connaît *à priori* C_m^1, on en conclut facilement $C_m^2, C_m^3, \ldots$.

La méthode en question se réduit donc : 1° à trouver la fonction dont le terme général ait pour coefficient la quantité cherchée; 2° cette fonction étant connue, à former son développement. La fonction en question est ce que l'on appelle, d'après Laplace, la *fonction génératrice* de l'ex-

pression à calculer. Laplace avait créé, à ce propos, un algorithme spécial, dont il a fait un fréquent usage dans sa *Théorie analytique des probabilités;* mais le calcul des fonctions génératrices est aujourd'hui à peu près oublié, parce que le calcul des résidus de Cauchy conduit plus sûrement au même but.

Lorsqu'il s'agit de trouver le terme général du développement de $f(x)$ ou, ce qui revient au même, de $f(e^{t\sqrt{-1}})$ suivant les puissances de $e^{t\sqrt{-1}}$, ce développement étant supposé possible *à priori*, on peut y parvenir comme il suit. Posons

$$f(e^{t\sqrt{-1}}) = a_0 + a_1 e^{t\sqrt{-1}} + \ldots + a_n e^{nt\sqrt{-1}} + \ldots$$
$$+ b_1 e^{-t\sqrt{-1}} + \ldots + b_n e^{-nt\sqrt{-1}} + \ldots;$$

si l'on multiplie les deux membres de cette formule par $\frac{e^{-nt\sqrt{-1}}}{2\pi}dt$, et si l'on intègre entre des limites α et $\alpha + 2\pi$, tous les termes du second membre s'évanouiront à l'exception du terme en $e^{nt\sqrt{-1}}$. En effet,

$$\int_\alpha^{\alpha+2\pi} e^{pt\sqrt{-1}}dt = \left(\frac{e^{pt\sqrt{-1}}}{p\sqrt{-1}}\right)_\alpha^{\alpha+2\pi} = 0,$$

toutes les fois que p n'est pas nul; et, si p est nul, la fonction placée sous le signe $\int$ se réduit à dt; par suite, l'intégrale prend la valeur 2π; on a donc

$$\frac{1}{2\pi}\int_\alpha^{\alpha+2\pi} f(e^{-t\sqrt{-1}})e^{-nt}dt = a_n \quad (*).$$

(*) Plus généralement, si l'on a

$$f(z) = a_0 + a_1 z + a_2 z^2 + \ldots + a_n z^n + \ldots,$$

Si l'on observe que l'on a, en supposant m et n entiers,

$$(a)\quad\left\{\begin{aligned}&\int_\alpha^{\alpha+2\pi}\cos mx\cos nx\,dx=\left\{\begin{array}{lll}0 & \text{si} & m\gtrless n,\\ \pi & \text{si} & m=n;\end{array}\right.\\ &\int_\alpha^{\alpha+2\pi}\sin mx\sin nx\,dx=\left\{\begin{array}{lll}0 & \text{si} & m\gtrless n,\\ \pi & \text{si} & m=n;\end{array}\right.\\ &\int_0^{2\pi}\cos mx\sin nx\,dx=0,\end{aligned}\right.$$

on trouvera facilement le terme général du développement de $f(x)$, sous la forme

$$(b)\quad\left\{\begin{aligned}f(x)=\mathrm{A}_0+\mathrm{A}_1\cos x+\mathrm{A}_2\cos 2x+\ldots+\mathrm{A}_n\cos nx+\ldots\\ +\mathrm{B}_1\sin x+\mathrm{B}_2\sin 2x+\ldots+\mathrm{B}_n\sin nx+\ldots,\end{aligned}\right.$$

lorsque l'on saura *à priori* que ce développement est possible, en suivant une marche analogue à celle qui nous a conduit au développement de $f\left(e^{t\sqrt{-1}}\right)$. Si l'on multiplie, en effet, par $\cos nx$ les deux membres de (b), et si l'on intègre entre les limites α et $\alpha+2\pi$, il viendra, en vertu de (a),

$$\int_\alpha^{\alpha+2\pi}f(x)\cos nx\,dx=\pi\mathrm{A}_n,$$

d'où

$$\mathrm{A}_n=\frac{1}{\pi}\int_\alpha^{\alpha+2\pi}f(x)\cos nx\,dx.$$

---

on sait que

$$a_n=\frac{1}{2\pi\sqrt{-1}}\int\frac{f(z)\,dz}{z^{n+1}}.$$

L'intégrale étant prise le long d'un contour contenant l'origine dans son intérieur, quand ce contour se réduit à un cercle dont le rayon est l'unité, on retrouve le résultat précédent.

Il y a exception pour le premier coefficient A_0, et l'on a, en intégrant de α à $\alpha + 2\pi$,

$$A_0 = \frac{1}{2\pi}\int_\alpha^{\alpha+2\pi} f(x)\,dx.$$

on trouve, d'une façon tout à fait analogue,

$$B_n = \frac{1}{\pi}\int_\alpha^{\alpha+2\pi} f(x)\sin nx\,dx$$

et, par conséquent,

$$f(x) = \frac{1}{2\pi}\int f(x)\,dx + \sum_1^\infty \cos nx \int f(x)\cos nx\,dx$$
$$+ \sum_1^\infty \sin nx \int f(x)\sin nx\,dx,$$

toutes les intégrales étant prises entre les limites α et $\alpha + 2\pi$.

Nous terminerons ces considérations par un beau théorème dû à Parseval. Si l'on a

$$\varphi\left(e^{t\sqrt{-1}}\right) = a_0 + a_1 e^{t\sqrt{-1}} + a_2 e^{2t\sqrt{-1}} + \ldots,$$
$$\psi\left(e^{-t\sqrt{-1}}\right) = b_0 + b_1 e^{-t\sqrt{-1}} + b_2 e^{-2t\sqrt{-1}} + \ldots,$$

en faisant le produit de ces deux formules et en intégrant, depuis $t = \alpha$ jusqu'à $t = \alpha + 2\pi$, on trouve

$$(\text{A})\ \frac{1}{2\pi}\int_\alpha^{\alpha+2\pi} \varphi\left(e^{t\sqrt{-1}}\right)\psi\left(e^{-t\sqrt{-1}}\right)dt = a_0 b_0 + a_1 b_1 + a_2 b_2 + \ldots.$$

Cette formule permet, comme on le voit, de trouver la valeur de la série dont le terme général est $a_n\, b_n$, quand on

connaît les valeurs des séries (supposées convergentes, lorsque le module de x est 1), dont les termes généraux sont $a_n x^n$ et $b_n x^n$; et c'est en cela que consiste le théorème de Parseval (*).

Évaluation approchée de P_m, C_m^n et A_m^n lorsque m et n sont très-grands.

Lorsque l'on veut évaluer l'un des nombres $P_m = 1.2\ldots m$, C_m^n ou A_m^n, on se trouve souvent arrêté par la longueur des calculs provenant de ce que m et n sont de grands nombres; Stirling a fait connaître une formule approchée, bien suffisante dans la plupart des cas, et qui ramène le calcul de P_m, C_m^n, A_m^n à celui de l'expression très-simple $n^n\sqrt{2\pi n}e^{-n}$, qui est à peu près égale à $1.2.3\ldots n$. La formule de Stirling, dont la démonstration a été rendue rigoureuse dans ces derniers temps par Cauchy, a été l'objet de travaux remarquables, parmi lesquels nous citerons ceux de Laplace, de Poisson, de Cauchy, de MM. Serret, Liouville, Binet, etc. Ces géomètres sont, en général, partis de la formule

$$1.2.3\ldots n = \int_0^\infty x^n e^{-x} dx.$$

La méthode que nous allons indiquer nous a paru à la fois la plus simple et la plus rigoureuse; elle a, du reste, l'avantage d'être une application immédiate des principes

(*) On pourrait aussi énoncer le théorème de Parseval en disant que si les séries, dont les termes généraux sont $a_n x^n$ et $b_n x^{-n}$, sont convergentes le long d'un cercle de rayon R et à l'intérieur de ce cercle, et ont pour valeurs respectives $\varphi(x)$ et $\psi(x)$, la série dont le terme général est $a_n b_n$ a pour valeur $\mathcal{E}\dfrac{\varphi(x)\psi(x)}{x}$, le résidu étant pris à l'intérieur du cercle en question.

que nous venons d'exposer; on a

$$e^{nx} = 1 + \frac{nx}{1} + \frac{n^2x^2}{1.2} + \ldots + \frac{n^nx^n}{1.2.3\ldots n} + \ldots.$$

Changeant x en $\cos t + \sqrt{-1}\sin t$, ou en $e^{t\sqrt{-1}}$, il vient

$$e^{n(\cos t + \sqrt{-1}\sin t)} = 1 + \frac{n}{1}e^{t\sqrt{-1}} + \frac{n^2 e^{2t\sqrt{-1}}}{1.2} + \ldots + \frac{n^n e^{nt\sqrt{-1}}}{1.2.3\ldots n} + \ldots.$$

Conformément à l'esprit de la méthode générale donnée plus haut, nous multiplierons par $\frac{1}{2\pi}e^{-nt\sqrt{-1}}dt$ les deux membres de cette formule et nous intégrerons de $-\pi$ à $+\pi$. Nous aurons alors

$$\frac{n^n}{1.2.3\ldots n} = \frac{1}{2\pi}\int_{-\pi}^{+\pi} e^{n(\cos t + \sqrt{-1}\sin t) - nt\sqrt{-1}}\,dt,$$

ou, ce qui revient au même, en observant que les éléments réels de l'intégrale sont égaux deux à deux et que les éléments imaginaires s'entre-détruisent,

$$\frac{n^n}{1.2.3\ldots n} = \frac{1}{\pi}\int_0^{\pi} e^{n\cos t}\cos n(t - \sin t)\,dt.$$

On peut observer que $\cos t$ est égal à $1 - 2\sin^2\frac{1}{2}t$; si l'on pose alors

$$\sin\frac{1}{2}t = \theta,$$

on trouve

$$(1)\quad \frac{n^n}{1.2.3\ldots n} = 2\frac{e^n}{\pi}\int_0^1 e^{-2\theta^2 n}\cos 2n\left(\arcsin\theta - \theta\sqrt{1-\theta^2}\right)\frac{d\theta}{\sqrt{1-\theta^2}};$$

or on a

$$\arcsin\theta = \theta + \frac{\theta^3}{3}\frac{1}{2} + \frac{\theta^5}{5}\frac{1.3}{2.4} + \ldots,$$

$$\theta\sqrt{1-\theta^2} = \theta - \theta^3\frac{1}{2} - \frac{1}{2.4}\theta^5 - \ldots.$$

On peut donc écrire, au lieu de l'équation (1),

$$\frac{n^n e^{-n}}{1.2.3\ldots n} = \frac{2}{\pi}\int_0^1 e^{-2n\theta^2}\cos 2n\left(\frac{2}{3}\theta^3+\omega\right)\frac{d\theta}{\sqrt{1-\theta^2}},$$

ω désignant un ensemble de termes contenant θ^5, au moins en facteur. Du reste, la série ω est toujours convergente quand θ varie de 0 à 1, excepté pour $x = 1$; mais, négligeant le dernier élément de l'intégrale, on ne néglige qu'une quantité infiniment petite, en sorte qu'il n'est pas nécessaire d'avoir égard à la divergence de la série dans ce cas extrême. Si maintenant on pose

$$\theta\sqrt{2n} = x,$$

on trouve

$$(2)\quad \frac{n^n e^{-n}\sqrt{2n}}{1.2.3\ldots n} = \frac{2}{\pi}\int_0^{\sqrt{2n}} e^{-x^2}\cos\left(\frac{2}{3}x^3\frac{1}{\sqrt{2n}}+\omega\right)\frac{dx}{\sqrt{1-\frac{x^2}{2n}}};$$

que l'on fasse maintenant $n = \infty$, le second membre de cette formule tendra vers la valeur de l'intégrale

$$(3)\qquad \frac{2}{\pi}\int_0^\infty e^{-x^2}dx.$$

Pour établir cette proposition en toute rigueur, on décomposera l'intégrale de la formule (2) en deux autres : la première aura pour limites zéro et $\sqrt[3]{2n}$ par exemple, la se-

conde aura pour limites $\sqrt[9]{2n}$ et $\sqrt{2n}$; il est facile de voir que cette dernière est nulle; en effet, elle est moindre que

$$\int_{\sqrt[9]{2n}}^{\sqrt{2n}} e^{-(2n)^{\frac{2}{9}}} \frac{dx}{\sqrt{1-\frac{x^2}{2n}}} \quad \text{ou} \quad e^{-(2n)^{\frac{2}{9}}}\sqrt{2n}\left(\frac{\pi}{2}-\arcsin\frac{\sqrt[9]{2n}}{\sqrt{2n}}\right);$$

mais le produit $e^{-(2n)^{\frac{2}{9}}}\sqrt{2n}$ a pour limite zéro quand n est infini; il en résulte que l'on peut écrire, au lieu de (2),

$$\frac{n^n e^{-n}\sqrt{2n}}{1.2.3\ldots n} = \frac{2}{\pi}\int_0^{\sqrt[9]{2n}} e^{-x^2}\cos\left(\frac{2}{3}x^2\frac{1}{\sqrt{2n}}+\omega\right)\frac{dx}{\sqrt{1-\frac{x^2}{2n}}}.$$

Cette fois on peut affirmer que le facteur de $e^{-x^2}dx$, sous le signe $\int$, reste toujours de la forme $1+\varepsilon$, en désignant par ε une quantité qui, pour $n=\infty$, s'évanouit; on a donc

$$\lim \frac{n^n e^{-n}\sqrt{2n}}{1.2.3\ldots n} = \frac{2}{\pi}\int_0^{\infty} e^{-x^2}dx \quad \text{pour} \quad n=\infty.$$

Si l'on remplace l'intégrale bien connue qui figure dans cette équation par sa valeur $\frac{1}{2}\sqrt{\pi}$, on a enfin

$$(1) \qquad \lim \frac{n^n e^{-n}\sqrt{2\pi n}}{1.2.3\ldots n} = 1.$$

On voit donc que, si n est très-grand, $n^n e^{-n}\sqrt{2\pi n}$ est une valeur approchée de $1.2.3\ldots n$. Je vais maintenant, en suivant une analyse développée par M. Serret, montrer que cette valeur est extrêmement approchée, lors même que

n n'est pas très-grand; posons

$$(2) \qquad \frac{n^n e^{-n}\sqrt{2\pi n}}{1.2.3\ldots n} = \varphi(n);$$

il en résulte

$$\frac{\varphi(n)}{\varphi(n+1)} = (n+1)^{-n} n^n \sqrt{\frac{n}{n+1}}\, e,$$

ou, en prenant les logarithmes népériens,

$$\log\varphi(n) - \log\varphi(n+1) = -n\log\left(1+\frac{1}{n}\right) - \frac{1}{2}\log\left(1+\frac{1}{n}\right) + 1;$$

d'où, changeant n en $n+1, \ldots$,

$$\log\varphi(n+1) - \log\varphi(n+2)$$
$$= -(n+1)\log\left(1+\frac{1}{n+1}\right) - \frac{1}{2}\log\left(1+\frac{1}{n+1}\right) + 1,$$

$$\log\varphi(n+2) - \log\varphi(n+3)$$
$$= -(n+2)\log\left(1+\frac{1}{n+2}\right) - \frac{1}{2}\log\left(1+\frac{1}{n+2}\right) + 1,$$

....

En ajoutant ces formules et en observant que $\log\varphi(n+m)$ tend vers zéro pour $m=\infty$, en vertu de l'équation (1), il vient

$$\log\varphi(n) = -\lim\left[\sum_0^m (n+m)\log\left(1+\frac{1}{n+m}\right) + \frac{1}{2}\sum_0^m \log\left(1+\frac{1}{n+m}\right) - (m+1)\right],$$
$$m=\infty.$$

Si nous développons les logarithmes par la formule

connue

$$l(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \ldots,$$

il vient

$$\log\varphi(n) = \lim\left[\frac{1}{2}\sum_0^m \frac{1}{(n+m)} - \frac{1}{3}\sum_0^m \frac{1}{(n+m)^2} + \ldots\right.$$

$$\left. - \frac{1}{2}\sum_0^m \frac{1}{n+m} + \frac{1}{4}\sum_0^m \frac{1}{(n+m)^2} - \ldots\right],$$

c'est-à-dire

$$\log\varphi(n) = -\frac{1}{12}\sum_0^\infty \frac{1}{(n+m)^2} + \frac{1}{12}\sum_0^\infty \frac{1}{(n+m)^3} - \ldots$$

$$\pm \frac{p-1}{2p(p+1)}\sum_0^\infty \frac{1}{(n+m)^p} - \ldots;$$

on en conclut, en remplaçant $\varphi(n)$ par sa valeur (2),

$$1.2.3\ldots n = n^n e^{-n}\sqrt{2\pi n}\, e^{\varpi(n)},$$

ϖ désignant la quantité

$$(3)\qquad \varpi(n) = \frac{1}{12}\sum_0^\infty \frac{1}{(n+m)^2} - \frac{1}{12}\sum_0^\infty \frac{1}{(n+m)^3} + \ldots.$$

Cette quantité ϖ est inférieure à son premier terme, lequel est lui-même inférieur, comme l'on sait, à

$$\frac{1}{12}\left(\frac{1}{n} + \frac{1}{2n} + \frac{1}{4n} + \ldots\right) \quad \text{ou à} \quad \frac{1}{6n};$$

donc on pourra poser

$$(4)\qquad 1.2.3\ldots n = n^n e^{-n}\sqrt{2\pi n}\,(1+\varepsilon),$$

et la quantité ε sera inférieure à $\frac{1}{6n} + \frac{1}{2}\left(\frac{1}{6n}\right)^2 + \ldots$, de sorte que, si n est assez grand, l'erreur relative sera à peu près $\frac{1}{6n}$. Cette erreur est tout à fait insignifiante dans les applications; on peut du reste l'évaluer au moyen de la formule (3).

La formule (4) est surtout utile pour évaluer le nombre des combinaisons de m objets pris n à n quand m et n sont très-grands. On a en effet, à peu près,

$$C_m^n = \frac{m!}{(m-n)!\,n!} = \frac{m^m}{n^n(m-n)^{m-n}}\sqrt{\frac{m}{2\pi n(m-n)}}.$$

La quantité A_m^n s'évaluerait d'une façon analogue.

Des combinaisons dans lesquelles entrent des objets de même nature et de nature différente.

Problème I. — *Trouver le nombre de manières distinctes dont on peut ranger m objets dont α sont identiques à a.*

Les objets en question pourraient se ranger de $P_m = m!$ manières différentes s'ils étaient tous différents; mais comme α d'entre eux sont identiques à a, il est clair que si l'on considère l'un des groupes et que si l'on y permute les objets a, on obtiendra P_α ou $\alpha!$ résultats identiques; mettant de côté ce groupe et ceux que l'on en déduit par les permutations des objets a, on pourra répéter la même opération sur un second groupe, et ainsi de suite, jusqu'à ce qu'on les ait tous mis de côté; il est clair que l'on aura ainsi formé $\frac{m!}{\alpha!}$ systèmes de groupes identiques. Le nombre de

groupes distincts que l'on peut former avec m objets dont α sont identiques à a est donc $\frac{m!}{\alpha!}$.

PROBLÈME II. — *Trouver le nombre de manières distinctes dont on peut ranger m objets dont α sont identiques à a, dont β sont identiques à b, dont γ sont identiques à c, etc.*

Supposons d'abord tous les objets distincts, il y aura P_m manières de les ranger; nous avons vu que, si l'on en supposait α identiques à a, les P_m groupes, d'abord différents, se partageaient en $\frac{P_m}{P_\alpha}$ systèmes de groupes identiques; par suite, nous n'avons plus en définitive que $\frac{P_m}{P_\alpha}$ groupes distincts. Si nous supposons maintenant que β objets deviennent identiques à b, en permutant dans l'un des nouveaux groupes les objets b, on obtiendra P_β groupes, que l'on pourra mettre de côté, et faisant partie des $\frac{P_m}{P_\alpha}$ groupes que nous avons considérés comme seuls distincts. En opérant de la même façon sur l'un des groupes que l'on n'a pas mis de côté, on obtiendra encore P_β résultats identiques, et ainsi de suite; donc, en définitive, le nombre des résultats distincts, lorsque α objets sont identiques à a et que β objets sont identiques à b, est $\frac{P_m}{P_\alpha P_\beta}$ ou $\frac{m!}{\alpha!\,\beta!}$; en supposant maintenant que γ objets deviennent identiques à c, on trouvera, pour le nombre des résultats distincts, $\frac{P_m}{P_\alpha P_\beta P_\gamma}$ ou $\frac{m!}{\alpha!\,\beta!\,\gamma!}$, et ainsi de suite.

On peut encore arriver à ce résultat comme il suit : considérons le produit

$$(a+b+c+\ldots+l)(a'+b'+c'+\ldots+l')$$
$$\times(a''+b''+c''+\ldots+l'')\ldots,$$

dans lequel le nombre des facteurs est m; le terme général de ce produit se composera de α facteurs a, de β facteurs b; ..., de λ facteurs l, de telle sorte que l'on ait

$$\alpha + \beta + \gamma + \ldots + \lambda = m.$$

Si l'on suppose

$$a = a' = a'' \ldots, \quad b = b' = b'' = \ldots;$$

tous les termes où entrent α facteurs a, β facteurs b, etc., deviennent identiques et égaux à $a^\alpha b^\beta c^\gamma, \ldots$, qui se trouvera ainsi répété autant de fois que l'on peut, sur les lettres $a, b, \ldots, l$, en prendre m, et les arranger de manière que α soient identiques à a, β identiques à b, etc. Ainsi le coefficient de $a^\alpha b^\beta c^\gamma \ldots l^\lambda$ dans $(a + b + c + \ldots + l)^m$ est le nombre de manières dont on peut ranger m objets, dont α sont identiques à a, etc.; il ne reste plus qu'à trouver ce coefficient, dont la valeur est bien $\frac{m!}{\alpha!\,\beta!\ldots\lambda!}$, ainsi qu'il a été prouvé en Algèbre.

Lorsque $m, \alpha, \beta, \ldots$ sont de grands nombres, la formule de Stirling fera connaître le résultat avec une approximation suffisante, en la bornant à son premier terme pour l'évaluation de chacun des facteurs $\alpha!\,\beta!\ldots m!$

Problème III. — *Trouver le nombre de groupes différents que l'on peut former en prenant n lettres sur s lettres données, les s lettres se composant de α lettres a, β lettres b, γ lettres $c, \ldots, \lambda$ lettres l.*

Nous regardons comme identiques deux groupes contenant le même nombre de lettres a, le même nombre de lettres b, etc. Supposons d'abord les s lettres considérées distinctes, le nombre des groupes contenant n lettres sera C_s^n; supposons maintenant que α lettres deviennent égales à a,

les combinaisons distinctes seront d'abord les combinaisons qui ne contiennent pas a et qui sont les combinaisons de $s-\alpha$ lettres n à n au nombre de $C_{s-\alpha}^{n}$, puis les combinaisons contenant a une fois; ces combinaisons s'obtiennent en ajoutant la lettre a aux combinaisons de $s-\alpha$ lettres prises $n-1$ à $n-1$; elles sont au nombre de $C_{s-\alpha}^{n-1}$, etc.; le nombre des groupes distincts que l'on pourra former, si α lettres deviennent égales à a, sera donc

$$C_{s-\alpha}^{n} + C_{s-\alpha}^{n-1} + \ldots + C_{s-\alpha}^{n-\alpha},$$

en arrêtant cette suite dès que l'exposant de C devient inférieur à zéro (s'il y a lieu).

Si non-seulement α lettres deviennent égales à a, mais encore si β lettres deviennent égales à b, on trouvera le nombre des groupes distincts en observant que $C_{s-\alpha-\beta}^{n}$ est le nombre des groupes où n'entrent ni a ni b; que, en général, si $\theta_{i,j}$ indique le nombre des groupes où a entre i fois et b j fois, les groupes en question se composeront des combinaisons de $s-\alpha-\beta$ lettres prises $n-i-j$ à $n-i-j$, à la suite desquelles on écrit i lettres a et j lettres b; donc

$$\theta_{i,j} = C_{s-\alpha-\beta}^{n-i-j}.$$

Le nombre cherché sera donc

$$\Sigma\theta_{ij} = \Sigma C_{s-\alpha-\beta}^{n-i-j},$$

où l'on devra faire varier i et j de zéro à α et de zéro à β, en convenant de faire $C_p^q = 0$ toutes les fois que $q < 0$.

On voit sans peine comment on étendrait cette solution au cas général.

Introduction à la théorie des séries trigonométriques.

Daniel Bernoulli, dans l'étude du mouvement des cordes vibrantes, a été conduit à représenter certaines fonctions au moyen de séries procédant suivant les sinus et les cosinus des multiples d'un même arc. Euler, en s'occupant de la même question, a donné l'expression des coefficients de ces séries; il est juste de dire, toutefois, que d'Alembert avait déjà indiqué aux géomètres la manière de résoudre cette question en montrant comment on pouvait trouver les coefficients du développement de $(1 + \alpha \cos x + \alpha^2)^{-\frac{1}{2}}$, ordonné suivant les cosinus des multiples de x. Plus tard, Lagrange fit connaître une formule toute semblable à celle d'Euler, en généralisant sa méthode d'interpolation. Enfin Fourier, à propos de ses recherches sur la chaleur, donna une extension très-grande à la théorie des séries trigonométriques; mais c'est à Dirichlet que l'on doit les premiers raisonnements rigoureux qui aient fixé le sens des formules découvertes par les géomètres que nous venons de citer. Cauchy, en s'aidant du calcul des résidus, a donné des formules très-générales qui lui ont permis d'aller encore beaucoup plus loin dans cette voie. Nous allons ici présenter la méthode que Dirichlet a fait connaître dans le *Journal de Crelle*, t. IV, p. 166.

Désignons par $f(x)$ une fonction quelconque de x assujettie à être finie entre les limites a et b de x; nous la supposerons, de plus, positive et continue; du reste, sa dérivée pourra être infinie ou indéterminée, enfin elle pourra être discontinue aux limites a et b. Cherchons la valeur de l'intégrale

$$V = \int_a^b f(x) \frac{\sin \omega x}{\sin x} dx,$$

dans laquelle ω est un nombre positif qui croît indéfiniment. Cette intégrale se rencontre fréquemment, comme nous le verrons dans le calcul des probabilités.

Si l'on pose $\omega x = t$, il vient

$$V = \frac{1}{\omega}\int_{a\omega}^{b\omega} f\left(\frac{t}{\omega}\right)\frac{\sin t}{\sin\frac{t}{\omega}}\,dt,$$

et l'on peut décomposer l'intégrale en une série d'autres de la forme

$$A_i = \frac{1}{\omega}\int_{\omega a + 2\pi i}^{\omega a + 2\pi(i+1)} f\left(\frac{t}{\omega}\right)\frac{\sin t}{\sin\frac{t}{\omega}}\,dt.$$

La dernière de ces intégrales seule aura des limites dont la différence pourra ne pas être 2π; mais, si ω est infini, on pourra négliger cette dernière intégrale, pourvu que b ne soit pas un multiple de π, parce que le sinus qui est au dénominateur s'annulerait. Laissons, pour le moment, cette difficulté de côté. Si, de plus, $\sin x$ ne s'annule pas entre les limites a et b, $\sin\frac{t}{\omega}$ conservera le même signe entre les limites $a + 2\pi i$ et $a + 2\pi(i+1)$; alors il est clair qu'en remplaçant $\dfrac{f\left(\frac{t}{\omega}\right)}{\sin\frac{t}{\omega}}$ par son maximum $\frac{1}{2}$M dans les éléments positifs, et par son minimum $\frac{1}{2}m$ dans les éléments négatifs de l'intégrale A_i, cette intégrale ne pourra qu'augmenter; mais, dans cette hypothèse, elle se réduit à $M - m$ en valeur absolue; on a donc

$$A_i < \frac{M - m}{\omega}$$

et

$$\mathrm{V} \quad \text{ou} \quad \sum \mathrm{A}_i < \sum \frac{(\mathrm{M}-m)}{\omega},$$

ou, en désignant par μ la plus grande des quantités $\mathrm{M} - m$,

$$\mathrm{V} < \mu \sum \frac{1}{\omega},$$

$$\mathrm{V} < \mu(b-a);$$

or, la fonction $\frac{f(x)}{\sin x}$ étant continue, μ tend vers zéro quand ω croît : la limite de V est donc zéro pour $\omega = \infty$. Nous avons supposé $f(x)$ positif, mais le raisonnement s'applique encore au cas où $f(x)$ serait négatif. Si $f(x)$ pouvait passer du positif au négatif dans l'intervalle compris entre a et b, en décomposant l'intégrale V en plusieurs autres dans lesquelles $f(x)$ conserverait le même signe, on verrait que V est encore nul, pourvu toutefois que $f(x)$ ne passe pas une infinité de fois du positif au négatif dans l'intervalle en question.

Tout change lorsque $\sin x$ s'annule entre les limites a et b. Supposons d'abord $\sin a = 0$ et $\sin x$ différent de zéro entre a et b, l'intégrale V peut alors se décomposer en deux autres, l'une

$$\mathrm{V}' = \int_a^{a+\varepsilon} f(x) \frac{\sin \omega x}{\sin x} dx,$$

et l'autre V″ prise entre les limites $a + \varepsilon$ et b, qui est nulle, lors même que ε est très-voisin de a, ce que nous supposerons; alors on a $\mathrm{V} = \mathrm{V}'$, et la formule précédente peut s'écrire

$$\mathrm{V} = f(\xi) \int_a^{a+\varepsilon} \frac{\sin \omega x}{\sin x} dx,$$

ξ désignant un nombre compris entre a et $a+\varepsilon$. Du reste, a étant de la forme $k\pi$, k désignant un nombre entier, on peut poser $x = a + \theta$, et l'on a

$$V = f(\xi) \int_0^\varepsilon \frac{\sin\omega(a+\theta)}{\sin(\theta+a)}\, d\theta.$$

Si a est égal à zéro, on aura

$$V = f(\xi) \int_0^\varepsilon \frac{\sin\omega\theta}{\sin\theta}\, d\theta.$$

Cette formule aura encore lieu si ω est un nombre impair indéfiniment croissant; nous nous placerons dans cette dernière hypothèse qui se trouve réalisée dans la pratique. Or, la limite ε étant très-petite, $\sin\theta$ peut être remplacé par θ sous le signe $\int$; il suffit, pour l'exactitude, d'écrire

$$V = f(\xi)\left(\frac{\eta}{\sin\eta}\right) \int_0^\varepsilon \frac{\sin\omega\theta}{\theta}\, d\theta.$$

η désignant un arc compris entre zéro et ε, mais sans qu'il soit nécessaire de reprendre des raisonnements déjà faits tout à l'heure, il est facile de voir que

$$\int_\varepsilon^\infty \frac{\sin\omega\theta}{\theta}\, d\theta$$

est nul pour $\omega = \infty$; on peut donc écrire

$$V = \lim f(\xi)\left(\frac{\eta}{\sin\eta}\right) \int_0^\infty \frac{\sin\omega\theta}{\theta}\, d\theta,$$

et, si l'on observe que l'intégrale

$$\int_0^\infty \frac{\sin\omega\theta}{\theta}\, d\theta$$

est égale à $\frac{\pi}{2}$, quel que soit ω, on aura

$$V = \frac{\pi}{2} f(a).$$

Le résultat auquel nous venons d'arriver est applicable évidemment à la limite b.

On peut déduire de là les conséquences qui suivent :

1° Quel que soit ω, pourvu qu'il augmente indéfiniment,

$$V = \int_a^b f(x) \frac{\sin \omega x}{\sin x} dx$$

tend vers zéro si a et b ne comprennent pas entre eux de multiples de π. Si a est nul, on a

$$V = \frac{\pi}{2} f(0);$$

enfin, si a et b sont de signes contraires, mais sont moindres que π en valeur absolue, on a

$$V = \pi f(0);$$

en effet l'intégrale V peut se décomposer en deux autres ayant la limite commune zéro, et toutes deux égales à $\frac{\pi}{2} f(0)$.

2° Si ω désigne un nombre impair infini, V est nul toutes les fois que a et b ne comprennent pas entre eux de multiple de π; V est égal à $\frac{\pi}{2} f(a)$ si a est multiple de π, ou à $\frac{\pi}{2} f(b)$ si b est multiple de π, $b - a$ restant inférieur à π. V est égal à $\pi f(c)$ si a et b comprennent un, et un seul multiple de π, désigné par c. En effet l'intégrale V peut se partager en deux autres ayant pour limites c, et, par suite, ayant chacune pour valeur $\frac{\pi}{2} f(c)$; enfin, si a et b

comprennent entre eux les multiples $c_1, c_2, c_3, \ldots$ de π, on a

$$(1) \qquad V = \pi[f(c_1) + f(c_2) + f(c_3) + \ldots],$$

et il faut augmenter le second membre de $\frac{\pi}{2} f(a)$ ou de $\frac{\pi}{2} f(b)$ si a et b sont multiples de π.

Il est à remarquer que rien ne suppose dans nos raisonnements la fonction f continue; il suffit qu'elle ne soit pas *sans cesse* discontinue; si, par exemple, $f(x)$ était discontinue pour $x = c$, on pourrait écrire

$$V = \int_a^c f(x) \frac{\sin \omega x}{\sin x} dx + \int_c^b \ldots dx,$$

et l'on conclurait de là que, si c n'est pas multiple de π, la discontinuité de $f(x)$ ne modifie en rien la valeur donnée plus haut de V; si c est multiple de π et si l'on désigne par $f_1(c)$ et $f_2(c)$ les deux valeurs de $f(x)$ pour $x = c$, le terme $f(c)$ de la formule (1) devra être remplacé par

$$\frac{\pi}{2}[f_1(c) + f_2(c)],$$

de sorte que la formule (1) peut encore s'écrire

$$V = \frac{\pi}{2}[f(a) + f_1(c_1) + f_2(c_1) + f_1(c_2) + f_2(c_2) + \ldots + f(b)],$$

$f_1(c)$ et $f_2(c)$ devenant des symboles identiques à $f(c)$, si $f(x)$ est continu pour $x = c$, et $f(a)$, $f(b)$ devant être remplacés par zéro, si a et b ne sont pas multiples de π.

Enfin les conclusions précédentes s'appliquent encore au cas où $f(x)$ serait de la forme $\varphi(x) + \psi(x)\sqrt{-1}$, car elles s'appliquent séparément à $\varphi(x)$ et à $\psi(x)$.

Théorie des séries trigonométriques.

Nous avons vu, au § I de ce Chapitre, que, la fonction $f(x)$ étant développable en série procédant suivant les sinus et les cosinus des multiples de l'arc x, ce développement était de la forme suivante :

$$f(x) = \frac{1}{2\pi}\int f(x)dx$$
$$+\frac{\cos x}{\pi}\int f(x)\cos x dx + \frac{\cos 2x}{\pi}\int f(x)\cos 2x dx + \ldots$$
$$+\frac{\sin x}{\pi}\int f(x)\sin x dx + \frac{\sin 2x}{\pi}\int f(x)\sin 2x dx + \ldots,$$

les limites étant α et $2\pi+\alpha$ pour toutes les intégrales. Cette formule peut se transformer en une autre plus simple si l'on fait entrer sous le signe $\int$ les facteurs $\cos x$, $\cos 2x, \ldots$, $\sin x$, $\sin 2x$; mais pour cela il faut changer, pour éviter toute confusion, la lettre par rapport à laquelle on intègre; nous la remplacerons par ξ, et nous aurons

$$f(x) = \frac{1}{2\pi}\int f(\xi)d\xi + \frac{1}{\pi}\int f(\xi)\cos x\cos\xi d\xi + \ldots,$$
$$+\frac{1}{\pi}\int f(\xi)\sin x\sin\xi dx + \ldots$$

ou bien encore

$$f(x) = \frac{1}{2\pi}\int f(\xi)d\xi + \frac{1}{\pi}\int f(\xi)\cos(\xi - x)d\xi$$
$$+\frac{1}{\pi}\int f(\xi)\cos 2(\xi - x)d\xi + \ldots.$$

Mais jusqu'ici cette formule ne peut avoir de valeur que *si l'on sait d'avance que le développement de $f(x)$ est pos-*

sible sous la forme d'une série trigonométrique. On peut alors se proposer de chercher la valeur du second membre de la formule précédente lorsque $f(x)$ est une fonction quelconque, et l'on peut y parvenir à l'aide des considérations suivantes. Pour plus de généralité, nous supposerons aux intégrales des limites arbitraires a et b, et nous poserons

$$W_n = \frac{1}{2\pi}\int_a^b f(\xi)d\xi + \frac{1}{\pi}\int_a^b f(\xi)\cos(\xi - x)d\xi$$
$$+ \frac{1}{\pi}\int_a^b f(\xi)\cos 2(\xi - x)d\xi + \ldots + \frac{1}{\pi}\int_a^b f(\xi)\cos n(\xi - x)d\xi,$$

en limitant pour le moment la série à ses premiers termes, afin de n'avoir point à nous embarrasser des questions de convergence. Nous pouvons écrire comme il suit la formule précédente :

$$(1)\quad \begin{cases} W_n = \frac{1}{\pi}\int_a^b f(\xi)\left[\frac{1}{2} + \cos(\xi - x) + \cos 2(\xi - x) + \ldots \right. \\ \qquad\qquad \left. + \cos n(\xi - x)\right]d\xi; \end{cases}$$

la quantité écrite entre crochets peut être remplacée par

$$\frac{\sin(2n+1)\frac{\xi - x}{2}}{2\sin\frac{\xi - x}{2}},$$

ce qu'il est facile de prouver en observant que, en général,

$$\frac{1}{2} + \cos a + \cos 2a + \ldots + \cos na$$

est égal à

$$\frac{1}{2}\left(e^{-na\sqrt{-1}} + e^{-(n-1)a\sqrt{-1}} + \ldots + 1 + e^{a\sqrt{-1}} + e^{2a\sqrt{-1}} + \ldots + e^{na\sqrt{-1}}\right)$$

ou, en regardant cette expression comme une progression géométrique, à

$$\frac{1}{2}\frac{e^{(n+1)\alpha\sqrt{-1}}-e^{-n\alpha\sqrt{-1}}}{e^{\alpha\sqrt{-1}}-1}=\frac{1}{2}\frac{e^{\left(n+\frac{1}{2}\right)\alpha\sqrt{-1}}-e^{-\left(n+\frac{1}{2}\right)\alpha\sqrt{-1}}}{e^{\frac{1}{2}\alpha\sqrt{-1}}-e^{-\frac{1}{2}\alpha\sqrt{-1}}}$$

$$=\frac{1}{2}\frac{\sin\left(n+\frac{1}{2}\right)\alpha}{\sin\frac{1}{2}\alpha};$$

on pourra donc écrire l'équation (1) comme il suit :

$$W_n=\frac{1}{2\pi}\int_a^b f(\xi)\frac{\sin(2n+1)\frac{\xi-x}{2}}{\sin\frac{\xi-x}{2}}d\xi;$$

ou bien, en posant $\xi = x + 2t$,

$$W_n=\frac{1}{\pi}\int_{\frac{a-x}{2}}^{\frac{b-x}{2}} f(x+2t)\frac{\sin(2n+1)t}{\sin t}dt.$$

En vertu de ce qui a été dit au § précédent, si l'on fait $n=\infty$, la valeur de W_n s'obtiendra en ajoutant les valeurs que prend $f(x+2t)$; quand on y remplace t par les multiples de π, compris entre $\frac{a-x}{2}$ et $\frac{b-x}{2}$, et si $\frac{a-x}{2}$ ou $\frac{b-x}{2}$ sont des multiples de π, il faut diviser par 2 les résultats de la substitution relatifs à ces limites.

Supposons que $\frac{a-x}{2}$ et $\frac{b-x}{2}$ comprennent entre eux les multiples $c_1, c_2, \ldots, c_i$ de π, on aura

$$\text{(A)}\qquad \frac{a-x}{2}<c_1,\quad c_2,\ldots,\quad c_i<\frac{b-x}{2}$$

ou

(B) $a < 2c_1 + x, \quad 2c_2 + x, \ldots, \quad 2c_i + x < b,$

et, par suite, d'après le théorème que nous venons de rappeler,

$$\lim W_n = f(x + 2c_1) + f(x + 2c_2) + \ldots + f(x + 2c_i),$$

ou, plus généralement,

$$\lim W_n = \frac{1}{2}[f_2(x + 2a) + f_1(x + 2c_1) + f_2(x + 2c_1) + \ldots],$$

si $\frac{a - x}{2}$, $\frac{b - x}{2}$ sont multiples de π et si la fonction présente des discontinuités.

Supposons, en particulier,

$$a < x < b, \quad b - a \leq 2\pi.$$

Cette formule est un cas particulier de (B); c'est celui où $c_1 = 0$ et où il n'existe pas d'autres multiples de π entre $\frac{a - x}{2}$ et $\frac{b - x}{2}$; dans ce cas on a

$$\lim W_n = f(x),$$

si $f(x)$ est continu, et

$$\lim W_n = \frac{1}{2}[f_1(x) + f_2(x)],$$

si $f(x)$ est discontinu; si x est égal à a ou à b, on a seulement

$$\lim W_n = \frac{1}{2}f(a) \quad \text{ou} \quad \frac{1}{2}f(b);$$

si $b = a + 2\pi$, on a, pour $x = a$,

$$\lim W_n = \frac{1}{2}f(a) + \frac{1}{2}f(a + 2\pi).$$

Si l'on se reporte alors à la formule (1), en l'écrivant

$$W_n = \frac{1}{2\pi}\int_a^b f(\xi)d\xi + \frac{1}{\pi}\int_a^b f(\xi)\cos(\xi - x)d\xi + \ldots$$
$$+ \frac{1}{\pi}\int_a^b f(\xi)\cos n(\xi - x)d\xi,$$

et si l'on y fait $n = \infty$, il vient

$$(2)\quad f(x) = \frac{1}{2\pi}\int_a^b f(\xi)d\xi + \sum_{n=1}^{n=\infty}\frac{1}{\pi}\int_a^b f(\xi)\cos n(\xi - x)d\xi$$

pour toutes les valeurs de x, de a et de b satisfaisant à la relation

$$a < x < b, \quad b - a \leq 2\pi.$$

Si, au contraire, x est compris en dehors des limites a et b, la formule (2) cesse d'être exacte, et son premier membre est nul si aucune des quantités $2k\pi + x$ n'est comprise entre a et b. Si l'on suppose $a = -\pi$, $b = +\pi$ et x compris entre $-\pi$ et $+\pi$, on a

$$f(x) = \frac{1}{2\pi}\int_{-\pi}^{+\pi} f(\xi)d\xi + \sum_1^\infty \frac{1}{\pi}\int_{-\pi}^{+\pi} f(\xi)d\xi \cos n(\xi - x)$$
$$= \frac{1}{2\pi}\int_{-\pi}^{+\pi} f(\xi)d\xi + \sum_1^\infty \frac{\cos nx}{\pi}\int_{-\pi}^{+\pi} f(\xi)\cos n\xi d\xi$$
$$+ \sum_1^\infty \frac{\sin nx}{\pi}\int_{-\pi}^{+\pi} f(\xi)\sin n\xi d\xi$$
$$= \frac{1}{2\pi}\sum_{-\infty}^{+\infty}\int_{-\pi}^{+\pi} f(\xi)\cos n(\xi - n)d\xi$$
$$= \frac{1}{2\pi}\sum_{-\infty}^{+\infty}\int_{-\pi}^{+\pi} f(\xi)e^{n(\xi - x)\sqrt{-1}}d\xi, \ldots$$

Ces dernières formules semblent bien moins générales que l'équation (2); mais il est facile de montrer qu'elles sont, au fond, identiques, et que l'on peut remonter de ces dernières à l'équation (2); admettons que l'on ait, pour toutes les valeurs de x comprises entre $-\pi$ et $+\pi$,

$$(3) \qquad 2\pi f(x) = \sum_{-\infty}^{+\infty} \int_{-\pi}^{+\pi} f(\xi) \cos n(\xi - x) d\xi.$$

Si l'on pose

$$\xi = a + \tau, \quad x = a + t, \quad f(a + t) = \mathrm{F}(t),$$

on aura

$$2\pi \mathrm{F}(t) = \sum_{-\infty}^{+\infty} \int_{-a-\pi}^{-a+\pi} \mathrm{F}(\tau) \cos n(\tau - t) d\tau,$$

et cette formule a lieu quelle que soit la fonction F et pour les valeurs de t, comprises entre les limites $-a-\pi$ et $-a+\pi$ arbitraires, différant de 2π; $\mathrm{F}(t)$ étant arbitraire, nous pouvons définir cette fonction par la condition d'être égale à zéro, quand τ varie de b à $-\alpha+\pi$ et de $-\alpha-\pi$ à a, et égale à $f(\tau)$, quand τ varie de a à b; on a alors, en négligeant dans les intégrales les portions nulles,

$$2\pi f(t) = \sum_{-\infty}^{+\infty} \int_{a}^{b} f(\tau) \cos n(\tau - t) d\tau,$$

pour toutes les valeurs de t comprises entre a et b; ce qui est la formule (2). Nous terminerons cette discussion en faisant observer que le second membre de la formule (2), qui a pour valeur $f(x)$, quand x est compris entre a et b, représente naturellement une fonction périodique, dont la période est 2π; en sorte que le second membre de la for-

mule (2) est une série convergente, que nous venons d'apprendre à sommer.

Rappelons-nous que $f(x)$ peut être discontinu, mais ne doit pas être infini entre les limites a et b; mais elle doit avoir un nombre limité de maxima ou de discontinuités. Ainsi l'on peut prendre pour $f(x)$ une fonction égale à 1, quand x varie de zéro à π et à -1, quand x varie de $-\pi$ à zéro; la formule (3) donne alors

$$f(x) = \frac{1}{2\pi}\sum_{-\infty}^{+\infty}\left[-\int_{-\pi}^{0}\cos n(\xi - x)dx + \int_{0}^{\pi}\cos n(\xi - x)dx\right]$$

ou

$$f(x) = \frac{4}{\pi}\left(\frac{\sin x}{1} + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \ldots\right),$$

pour $x = 0$; la valeur du second membre doit être $\frac{-1+1}{2}$ ou zéro, ce qu'il est facile de constater, et pour $x = \pi$ on doit avoir $-\frac{1}{2} + \frac{1}{2}$, c'est-à-dire encore zéro; pour $x = \frac{\pi}{2}$ on doit avoir 1; donc

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \ldots,$$

ce que l'on savait déjà.

Ces formules s'étendent facilement au cas de plusieurs variables; leurs applications à toutes les branches des sciences mathématiques et physico-mathématiques sont innombrables; nous nous bornerons ici aux seuls développements qui intéressent l'analyse des chances.

Formule de Fourier.

Reprenons la formule (2) du paragraphe précédent et faisons $a = 0$, $b = 2\pi$, nous aurons

$$(1)\qquad f(x) = \frac{1}{2\pi}\sum_{-\infty}^{+\infty}\int_0^{2\pi} f(\xi)\cos n(\xi - x)\,d\xi,$$

et, en posant

$$\xi = \frac{2\pi(\tau - a)}{b - a}, \quad x = \frac{2\pi(t - a)}{b - a},$$

il vient

$$f\left(\frac{2\pi t - 2\pi a}{b - a}\right) = \frac{1}{2\pi}\sum_{-\infty}^{+\infty}\int_a^b f\left(2\pi\frac{\tau - a}{b - a}\right)\cos n\frac{2\pi(\tau - t)}{b - a}\frac{2\pi\,d\tau}{b - a};$$

en posant

$$f\left(\frac{2\pi t - 2\pi a}{b - a}\right) = \varphi(t)$$

et en réduisant, il vient

$$(2)\qquad \varphi(t) = \frac{1}{b - a}\sum_{-\infty}^{+\infty}\int_a^b \varphi(\tau)\cos\frac{2\pi n(\tau - t)}{b - a}\,d\tau,$$

et cette formule a lieu pour toutes les valeurs de t, comprises entre a et b, car l'équation (1) avait lieu pour les valeurs de x comprises entre zéro et 2π; et quand on suppose x compris entre zéro et π, il faut supposer t compris entre a et b.

Rien n'empêche maintenant de faire $a = -\infty$, $b = \infty$ et de poser

$$\frac{1}{b - a} = du, \quad \frac{n}{b - a} = u;$$

alors l'équation (2) s'écrit

$$(3) \qquad \varphi(t) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \varphi(\tau)\cos 2\pi u(\tau - t)\,d\tau\,du.$$

Et l'intégration, par rapport à τ, doit être effectuée la première, comme cela résulte de notre mode de démonstration. La formule (3) est celle que l'on appelle *la formule de Fourier ;* on peut lui donner la forme

$$\varphi(t) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \varphi(\tau) e^{2\pi u(\tau - t)\sqrt{-1}}\,d\tau\,du$$

ou bien encore

$$(4) \qquad \varphi(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \varphi(t) e^{u(\tau - t)\sqrt{-1}}\,d\tau\,du \quad (*).$$

(*) C'est ici l'occasion de rectifier une erreur généralement répandue parmi les physico-mathématiciens et chez les meilleurs esprits. On a cru pouvoir déduire de (4), en différentiant par rapport à t,

$$(5) \qquad \varphi'(t) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \varphi(\tau) e^{2\pi u(\tau - t)\sqrt{-1}}\,\frac{du\,d\tau}{\sqrt{-1}}\,2\pi u.$$

Mais cette formule est *ordinairement* fausse ; les cas où elle est exacte sont *exceptionnels :* je vais les signaler. D'abord je ferai observer que la formule (1) ne peut pas être différentiée par rapport à x, car elle perd sa convergence pour des valeurs imaginaires de x ; les cas où elle *ne perd pas* sa convergence sont *exceptionnels*. En effet, si la série qui constitue le second membre de (1) ne perd pas sa convergence pour des valeurs imaginaires de x, cette série représente une fonction continue monodrome et monogène de $e^{(\xi - x)\sqrt{-1}}$, c'est-à-dire une fonction périodique continue, monodrome et monogène de x. Donc $f(x)$ devrait être périodique pour que (1) eût lieu pour les valeurs imaginaires de x. Mais, en vertu d'un théorème que j'ai établi pour la première fois en 1863, on ne peut différentier une série que pour les valeurs de la variable pour lesquelles la série reste convergente, et *tout autour* desquelles elle reste convergente, ce qui n'est pas ici le cas. Or la formule (4) n'est qu'une transformée de (1) ; on ne pourra donc en dé-

Pour faire une application de la formule de Fourier, on peut supposer $\varphi(\tau) = e^{-\tau}$ pour les valeurs positives de τ et égale à e^{τ} pour les valeurs négatives de τ; on a alors

$$\varphi(t) = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{0} e^{\tau+2\pi u(\tau-t)\sqrt{-1}} d\tau + \int_{0}^{\infty} e^{-\tau+2\pi u(\tau-t)\sqrt{-1}} d\tau \right] du$$

ou bien

$$\varphi(t) = \int_{-\infty}^{+\infty} \int_{0}^{\infty} e^{-\tau} \left[e^{-2\pi u(t+\tau)\sqrt{-1}} + e^{-2\pi u(t-\tau)\sqrt{-1}} \right] du\, d\tau,$$

ou enfin

$$\varphi(t) = 2 \int_{-\infty}^{+\infty} \frac{e^{-2\pi t u\sqrt{-1}}}{1 + 4\pi^2 u^2} du,$$

ce que l'on peut encore écrire

$$\varphi(t) = 4 \int_{0}^{\infty} \frac{\cos 2\pi t u}{1 + 4\pi^2 u^2} du.$$

Si l'on pose $2\pi u = x$, on a

$$\varphi(t) = \frac{2}{\pi} \int_{0}^{\infty} \frac{\cos t x}{1 + x^2} dx.$$

Valeur de quelques intégrales définies.

Je terminerai cette Introduction en rappelant les valeurs

---

duire (5) que si $\varphi(t)$ est une fonction *très-particulière*. Du reste, j'ai déjà attiré dans une autre occasion l'attention des géomètres sur la différentiation, sous le signe $\int$, et je crois avoir montré le premier que l'on abusait de cette règle, soumise à des restrictions qui ne sont encore indiquées que dans mon *Traité des résidus* ou dans le *Traité des fonctions* de M. Casorati.

3

de quelques intégrales dont nous aurons souvent besoin :

(A) $$\int_0^\infty e^{-x^2}dx = \frac{1}{2}\int_{-\infty}^{+\infty} e^{-x^2}dx = \frac{\sqrt{\pi}}{2},$$

(B) $$\int_0^\infty e^{-a^2x^2}dx = \frac{\sqrt{\pi}}{2a},$$

(C) $$\int_0^\infty \cos 2bx\, e^{-a^2x^2}dx = \frac{\sqrt{\pi}}{2a} e^{-\frac{b^2}{a^2}},$$

(D) $$\int_0^\infty \frac{\sin 2bx}{x} e^{-a^2x^2}dx = \frac{\sqrt{\pi}}{a}\int_0^b e^{-\frac{b^2}{a^2}}db.$$

La formule (A) est classique, (B) s'en déduit par un changement de variables, (C) s'obtient en changeant, dans (B), x en $x + \frac{b}{a^2}\sqrt{-1}$; enfin (D) s'obtient en intégrant (C) par rapport à b.

L'intégrale

$$\frac{2}{\sqrt{\pi}}\int_0^\gamma e^{-x^2}dx = \Theta(\gamma)$$

est très-voisine de l'unité, même pour des valeurs de γ relativement peu considérables ; cela ressort de la formule (A) ; cette intégrale joue un rôle important dans le calcul des probabilités. Voici une Table contenant quelques-unes de ses valeurs les plus usuelles :

Valeurs de γ.	Valeurs de $\Theta(\gamma) = \frac{2}{\sqrt{\pi}}\int_0^\gamma e^{-x^2}dx$.
0,476 936 3	$\frac{1}{2}$
1	0,842 700 8
1,163 087 2	0,9
1,821 386 4	0,99
2,327 675 4	0,999
2,751 065 4	0,9999
3	0,99997
....	......
∞	1

CHAPITRE II.

EXPOSITION DES MÉTHODES.

Définitions.

Le calcul des probabilités a pour but la mesure des chances qui président à l'arrivée des événements soumis au hasard.

Bien que l'arrivée de certains événements ne puisse point être prévue avec certitude, il n'en est pas moins vrai que leur chance est susceptible de plus ou de moins, et il y a lieu de chercher une mesure précise de cette chance. Un exemple va servir à faire comprendre notre pensée : Considérons un jeu de cartes composé de huit trèfles, huit cœurs, huit piques et huit carreaux; si l'on tire une carte au hasard dans le jeu, il est évident que la chance de tirer un trèfle ou un cœur est la même, que la chance de tirer un trèfle est plus grande que celle de tirer l'as de cœur, etc. Ainsi il est bien clair, dans l'exemple que nous venons de choisir, que l'idée de grandeur vient se joindre à celle de chance. Essayons maintenant d'analyser l'opération qui se fait dans notre esprit à propos de l'évaluation vague de la chance dont nous venons de parler. Si nous estimons égales les chances de tirer trèfle et cœur, c'est que le nombre des trèfles est égal à celui des cœurs, et que le nombre des cas dans lesquels on peut tirer trèfle est égal à celui des cas où l'on peut tirer cœur.

Au contraire, sur trente-deux cartes qui peuvent se pré-

senter avec la même facilité sous la main, quatre seront des as et huit seront des trèfles, et il est *naturel* de dire qu'il y a juste deux fois plus de chance pour tirer un trèfle que pour tirer un as; il semble donc *naturel* aussi de mesurer les chances d'arrivée des événements au moyen du nombre des cas favorables à l'arrivée de ces événements, mais on comprend que pour cela il est indispensable que les événements dont on compare les chances, se rapportent à un même nombre de cas favorables ou non à leurs arrivées respectives; je m'explique : supposons que du jeu de cartes considéré tout d'abord on retranche toutes les figures, la chance de tirer un as sera plus grande. Cependant le nombre des cas favorables est resté quatre, c'est toujours le nombre des as restés dans le jeu; ce qui a varié c'est le nombre des cas possibles qui a diminué de douze, c'est-à-dire du nombre des cartes enlevées. Ce que nous venons de dire justifiera suffisamment la définition que nous allons poser :

La PROBABILITÉ *d'un événement dû au hasard est le rapport du nombre de cas favorables à l'arrivée de cet événement au nombre total des cas qui peuvent se présenter quand on attend cet événement, tous ces cas, favorables ou non, étant censés* ÉGALEMENT POSSIBLES.

Par exemple, dans un jeu de cartes, la probabilité de tirer un trèfle est $\frac{1}{4}$, parce que l'événement attendu (la sortie de trèfle) peut se présenter de huit manières, qui sont représentées par les huit trèfles contenus dans le jeu, et que l'on peut du reste tirer l'une quelconque des trente-deux cartes du jeu avec la même facilité : ces trente-deux cartes représentent autant de cas possibles et également possibles; la probabilité en question est donc $\frac{8}{32}$ ou $\frac{1}{4}$, comme nous l'avions annoncé.

Une urne contient deux boules blanches et trois boules noires, en tout cinq; la probabilité d'extraire une boule blanche en tirant au hasard dans l'urne sera $\frac{2}{5}$, parce que le nombre des cas favorables est 2 et que celui des cas possibles et également possibles est 5.

La définition que nous venons de donner de la probabilité semble exclure l'estimation mathématique des chances de certains événements, ceux dont l'arrivée ne se rapporte pas à un nombre de cas également possibles; mais un examen attentif des questions permet presque toujours de décomposer les cas inégalement possibles en d'autres plus simples, et dont les arrivées ont des chances identiques.

Reprenons, par exemple, l'urne contenant deux boules blanches et trois boules noires; supposons les boules noires sphériques et égales, et les boules blanches formées de l'accouplement de deux sphères égales aux boules noires et reliées par un fil, il est clair que la *facilité* de l'extraction d'une boule blanche sera plus grande que celle d'une boule noire, et il serait absurde de dire que la probabilité de l'extraction d'une boule blanche est $\frac{2}{5}$. Mais on peut considérer chaque boule blanche comme formée de deux parties, sur lesquelles on peut mettre la main avec la même facilité que sur une boule noire; alors le nombre des cas favorables devient 4, celui des cas possibles 7, et la probabilité cherchée $\frac{4}{7}$.

Il est important d'observer que, quel que soit le mode de subdivision des cas possibles et également possibles, la probabilité d'un événement conserve la même valeur; cette proposition est l'axiome fondamental du calcul des probabilités.

De la définition que nous avons donnée de la probabilité d'un événement, il résulte que toute probabilité est un nombre compris entre zéro et l'unité; car le nombre des cas favorables à l'arrivée d'un événement E ne peut surpasser le nombre m des cas possibles; on a donc

$$o \leq n \leq m,$$

d'où l'on tire

$$o \leq \frac{n}{m} \leq 1;$$

$\frac{n}{m}$ est par définition la probabilité de E; cette probabilité est donc nécessairement comprise entre zéro et 1. Elle peut être zéro, et, dans ce cas, l'événement dont on cherche la probabilité, n'ayant aucune chance favorable, ne peut se présenter; elle peut être 1, et alors tous les cas possibles sont favorables, par suite l'événement attendu est certain. Ainsi zéro et 1 sont des symboles de certitude.

Nous devons cependant apporter une restriction aux dernières propositions que nous venons d'énoncer : nous avons supposé les cas possibles et favorables en nombre limité; plus loin nous verrons qu'il y a lieu de considérer quelquefois des nombres illimités de cas, et alors zéro et 1 ne sont plus des symboles de certitude *mathématique*, mais de certitude *morale*.

Recherche de la probabilité par les méthodes directes

PROBLÈME I. — *Une urne contient α boules blanches, β boules noires, γ boules rouges, etc.; on demande la probabilité d'extraire une boule blanche, une boule noire, une boule rouge, etc.*

La probabilité d'extraire une boule blanche s'obtiendra

en observant que le nombre des cas qui peuvent se présenter est égal à celui de toutes les boules ou $\alpha+\beta+\gamma+\ldots$; le nombre des cas favorables est celui des boules blanches ou α. La probabilité d'extraire une boule blanche est donc

$$\frac{\alpha}{\alpha+\beta+\gamma+\ldots};$$

celle d'extraire une boule noire serait

$$\frac{\beta}{\alpha+\beta+\gamma+\ldots},\ldots.$$

PROBLÈME II. — *Une urne contient α boules blanches et β boules noires; on demande la probabilité d'extraire d'un seul coup a boules blanches sans boules noires.*

Les cas qui peuvent se présenter, et avec des chances égales, sont ceux où l'on prend a boules sur $\alpha+\beta$ que contient l'urne, et cela sans avoir égard à l'ordre où on les prend; le nombre de ces cas est $C^a_{\alpha+\beta}$. Les cas favorables sont ceux dans lesquels on prend a boules blanches; leur nombre est celui des combinaisons de α objets pris a à a ou C^a_α. La probabilité cherchée est donc

$$C^a_\alpha : C^a_{\alpha+\beta} \quad \text{ou} \quad \frac{\alpha(\alpha-1)\ldots(\alpha-a+1)}{(\alpha+\beta)(\alpha+\beta-1)\ldots(\alpha+\beta-a+1)}.$$

PROBLÈME III. — *Une urne contient α boules blanches et β boules noires; on demande la probabilité d'extraire d'un seul coup a boules blanches et b boules noires.*

Le nombre des cas possibles est celui des combinaisons de $\alpha+\beta$ objets $a+b$ à $a+b$, ainsi qu'il est facile de le prouver en raisonnant comme tout à l'heure. Les cas favorables sont ceux où l'on tire a boules blanches et b boules

noires; or on peut tirer a boules blanches de C_α^a manières, comme cela a été prouvé dans l'exemple qui précède; on peut tirer b boules noires de C_β^b manières; enfin on peut tirer à la fois a boules blanches et b boules noires de $C_\alpha^a . C_\beta^b$ manières; car à chacune des manières d'extraire a boules blanches correspondent C_β^b manières d'extraire b boules noires, ce qui fait bien en tout $C_\alpha^a . C_\beta^b$ cas favorables distincts. La probabilité cherchée est donc finalement

$$\frac{C_\alpha^a . C_\beta^b}{C_{\alpha+\beta}^{a+b}}.$$

PROBLÈME IV. — *La loterie royale établie sous l'ancien régime, supprimée en 1793, rétablie en 1797 et supprimée définitivement en 1839, se composait de quatre-vingt-dix numéros dont on en tirait cinq au sort; on pouvait désigner un, deux, trois ou quatre numéros à l'avance: c'était ce que l'on appelait jouer l'extrait, l'ambe, le terne, le quaterne; on a souvent parlé du quine, mais il ne se jouait pas; enfin la mise des joueurs sur le quaterne ne pouvait pas dépasser 12 francs.*

Le joueur gagnant l'extrait recevait 15 fois sa mise, celui qui gagnait l'ambe recevait 270 fois sa mise, le terne rapportait 5500 fois la mise et le quaterne 7500 fois. Cherchons dans ces conditions quelles étaient les probabilités que les joueurs avaient de gagner l'extrait, l'ambe, le terne, le quaterne, le quine.

Cherchons d'abord la probabilité de l'extrait simple: soit n le nombre total des numéros et m le nombre des numéros tirés, le nombre des cas possibles est évidemment C_n^m, car on peut tirer avec une égale facilité l'une quelconque des combinaisons de n numéros pris m à m. Le nombre des cas favorables est celui des combinaisons de n objets pris m à m,

dans lesquelles un des objets désigné à l'avance se trouve compris. Otons cet objet, on aura les combinaisons de $n-1$ objets restants $m-1$ à $m-1$; le nombre des cas favorables est donc C_{n-1}^{m-1}, et, par suite, la probabilité de l'extrait est $C_{n-1}^{m-1} : C_n^m$ pour $n = 90$, $m = 5$; on a

$$\frac{89.88.87.86}{1.2.3.4} : \frac{90.89.88.87.86}{1.2.3.4.5} = \frac{5}{90} = \frac{1}{18}.$$

L'État aurait donc dû donner au joueur 18 fois sa mise et non 15 fois (*).

Cherchons maintenant la probabilité de tirer α numéros désignés à l'avance.

Le nombre des cas possibles reste comme tout à l'heure C_n^m; pour avoir le nombre des cas favorables, observons que ces cas sont les combinaisons de n numéros pris m à m, dans lesquelles entrent α numéros désignés; supprimons ces numéros, nous aurons des combinaisons de $n-\alpha$ objets pris $m-\alpha$ à $m-\alpha$, nous aurons même toutes ces combinaisons; car, en ajoutant aux combinaisons des $n-\alpha$ numéros pris $m-\alpha$ à $m-\alpha$ les numéros ôtés d'abord, on retrouvera des résultats qui seront des combinaisons de m numéros pris n à n. Enfin il va sans dire que les résultats considérés seront distincts; car, s'ils ne l'étaient pas, en y remettant les lettres ôtées, on obtiendrait des cas favorables qui ne seraient pas distincts non plus. En résumé, le nombre des cas favorables est $C_{n-\alpha}^{m-\alpha}$, et, par suite, la probabilité cherchée est $C_{n-\alpha}^{m-\alpha} : C_n^m$.

En faisant successivement $\alpha = 1, 2, 3, 4, 5$ et $n = 90$,

(*) Nous verrons plus loin que dans un jeu équitable les mises doivent être en raison directe des probabilités que les joueurs ont de gagner; mais on peut déjà regarder ce principe comme indiqué par les règles du bon sens.



$m = 5$, on a

La probabilité de l'extrait simple...	$\frac{1}{18}$;
» de l'ambe..........	$\frac{2}{801}$;
» du terne..........	$\frac{1}{11748}$;
» du quaterne........	$\frac{1}{511038}$;
» du quine	$\frac{1}{43949268}$.

Ces nombres, comme on voit, n'étaient pas en rapport avec les gains que les joueurs pouvaient espérer.

Problème V. — *Quelle est la probabilité de tirer, dans une loterie de n numéros, α numéros désignés sur m, mais à des places déterminées à l'avance?*

Pour résoudre cette question, nous regarderons non plus chaque combinaison de n numéros m à m, mais chaque arrangement de n numéros m à m comme un cas favorable ou possible, car cette fois il faut tenir compte de l'ordre dans lequel ils se présentent. Le nombre des cas possibles sera alors A_n^m, celui des cas favorables sera celui des arrangements de n numéros m à m, dans lesquels entrent α numéros donnés dans un ordre donné. Supprimons ces α numéros, nous obtiendrons des arrangements de $n-\alpha$ numéros $m-\alpha$ à $m-\alpha$, tous distincts, car, s'ils n'étaient pas distincts, en intercalant à leurs places les numéros désignés, on n'aurait pas des résultats distincts; enfin il est clair que tous les arrangements de $n-\alpha$ numéros $m-\alpha$ à $m-\alpha$ s'y trouveront. La probabilité demandée est donc

$$A_{n-\alpha}^{m-\alpha} : A_n^m.$$

Dans l'ancienne loterie, on pouvait ainsi placer un enjeu sur des numéros dont la place était désignée à l'avance : on jouait alors l'extrait déterminé, l'ambe déterminé, le terne déterminé, etc. La probabilité de l'extrait déterminé était

$$\frac{89.88.87.86}{90.89.88.87.86} = \frac{1}{90},$$

c'est-à-dire cinq fois plus faible que celle de l'extrait simple, ce qui était à peu près évident *à priori*, si l'on réfléchit que les cinq extraits déterminés ayant *à priori* la même probabilité, l'un quelconque d'entre eux se présentera cinq fois quand l'un d'eux, déterminé à l'avance, se présentera une fois seulement. Ce raisonnement sera rendu rigoureux un peu plus loin.

Probabilité composée.

THÉORÈME I. — *Lorsqu'un événement* E *se compose du concours de plusieurs autres* $e_1, e_2, e_3, \ldots$, *dont les arrivées ne se gênent en aucune façon, la probabilité de* E *est le produit des probabilités de* $e_1, e_2, \ldots$.

Soit, en général, m_i le nombre des cas favorables à l'arrivée de e_i et n_i le nombre des cas possibles, la probabilité de e_i sera $\frac{m_i}{n_i}$. Considérons d'abord l'événement qui se compose du concours de e_1 et de e_2 ; si ces événements ne se gênent pas, en d'autres termes s'ils sont tout à fait indépendants, à chacun des m_1 cas favorables à l'arrivée de e_1 correspondront m_2 cas favorables à l'arrivée de e_2 ; donc, quand on attend l'événement $e_1 e_2$, le nombre des cas favorables est $m_1 m_2$; on verrait de même que le nombre des cas

possibles est $n_1 n_2$; donc la probabilité du concours de e_1 et de e_2 est

$$\frac{m_1}{n_1}\frac{m_2}{n_2},$$

or $\frac{m_1}{n_1}$ et $\frac{m_2}{n_2}$ sont les probabilités de e_1 et de e_2; le théorème est donc démontré pour le cas de deux événements; si l'on considère alors l'événement composé de e_1, de e_2 et de e_3, on pourra le considérer comme composé du concours de $e_1 e_2$, d'une part, dont la probabilité est $\frac{m_1}{n_1}\frac{m_2}{n_2}$, et de e_3 dont la probabilité est $\frac{m_3}{n_3}$; sa probabilité sera, d'après ce que nous venons de voir,

$$\frac{m_1}{n_1}\frac{m_2}{n_2}\times\frac{m_3}{n_3}=\frac{m_1}{n_1}\frac{m_2}{n_2}\frac{m_3}{n_3},$$

et ainsi de suite pour un nombre quelconque d'événements $e_1\, e_2\, e_3 \ldots.$

Théorème II. — *Lorsqu'un événement* F *dépend du concours de deux autres* E *et* E′, *et que l'arrivée de* E′ *est subordonnée à celle de* E, *la probabilité de* F *est le produit de la probabilité de* E, *par la probabilité que* E *ayant eu lieu* E′ *aura lieu.*

En effet, désignons par p le nombre total des cas possibles, et également possibles, qui peuvent se présenter quand on attend l'arrivée de l'événement F qui dépend du concours de E et de E′. Sur ces p cas, supposons qu'il y en ait f favorables à l'arrivée de l'événement E. La probabilité de E sera alors $\frac{f}{p}$; mais quand l'événement E est arrivé, l'événement E′ peut avoir lieu d'un certain nombre de manières que l'on peut désigner par f'. Ainsi f' est le nombre

des cas favorables à l'arrivée de cet événement E′ quand E a déjà eu lieu; mais E pouvant avoir lieu de f manières, f est le nombre des cas possibles quand on attend l'arrivée de E′. Ainsi $\frac{f'}{f}$ sera la probabilité de E′ quand E aura eu lieu; mais la probabilité de l'événement F est $\frac{f'}{p}$, car p est le nombre des cas possibles et f' est le nombre des cas favorables à l'arrivée successive de E et de E′, c'est-à-dire de F; or $\frac{f'}{p} = \frac{f}{p}\,\frac{f'}{f}$: le théorème est donc démontré.

C'est dans ces théorèmes que consiste ce que l'on appelle *le principe de la probabilité composée*. Pour bien les faire comprendre, nous allons les appliquer à quelques exemples.

Problème I. — *Étant données deux urnes contenant l'une a boules blanches et b noires, l'autre a′ blanches et b′ noires, on demande la probabilité de tirer une boule blanche de chaque urne.*

L'événement dont on demande la probabilité dépend du concours de deux événements dont les arrivées ne se gênent pas, et, par conséquent, sa probabilité s'obtiendra en multipliant la probabilité d'extraire une blanche de la première urne par la probabilité d'extraire une blanche de la seconde : on trouve ainsi, pour le nombre demandé,

$$\frac{a}{a+b}\,\frac{a'}{a'+b'},$$

résultat auquel il serait facile d'arriver directement.

Problème II. — *Une urne contient a boules blanches et b noires; on tire successivement de cette urne n boules; on demande la probabilité que la première soit blanche,*

la seconde noire, la troisième noire, etc., les couleurs se succédant ainsi dans un ordre déterminé.

1° Supposons d'abord que l'on remette les boules dans l'urne après chaque tirage, la sortie successive de plusieurs boules est un événement composé, la probabilité que la première boule tirée sera blanche, la seconde noire, etc., sera le produit des probabilités de tirer une blanche, de tirer une noire, etc., et l'on aura, en définitive, pour expression de la probabilité cherchée,

$$\frac{a}{a+b}\,\frac{b}{a+b}\,\frac{b}{a+b}\cdots,$$

et, en général,

$$\left(\frac{a}{a+b}\right)^{\mu}\left(\frac{b}{a+b}\right)^{\nu};$$

μ désignant le nombre des blanches que l'on veut voir sortir et ν celui des noires; cette probabilité (et c'était évident *à priori*) ne dépend donc pas de l'ordre dans lequel on veut voir sortir les boules.

2° Supposons maintenant que les boules tirées ne soient pas remises dans l'urne, la probabilité de tirer une blanche, puis une noire, l'arrivée du second événement dépendant de celle du premier, sera égale au produit de la probabilité $\frac{a}{a+b}$ de tirer une blanche, par la probabilité qu'une blanche ayant été tirée on tire une noire au second coup. Cette dernière probabilité est $\frac{b}{a+b-1}$; en effet, le nombre des cas favorables est b, celui des cas possibles $a-1+b$; car il n'y a plus que $a-1$ boules blanches dans l'urne, une boule blanche ayant été tirée au premier coup. On voit ainsi que la probabilité de tirer successivement une

blanche et une noire est

$$\frac{a}{a+b}\,\frac{b}{a+b-1}.$$

Supposons que l'on demande la probabilité de tirer, en outre, une noire au troisième coup, il faudra multiplier la probabilité que nous venons de trouver par la probabilité de tirer une noire; en supposant que l'on ait ôté préalablement de l'urne une blanche et une noire, cette dernière probabilité est $\frac{b-1}{a+b-2}$ et, par suite, la probabilité demandée est

$$\frac{a.b.(b-1)}{(a+b)(a+b-1)(a+b-2)};$$

en continuant ce raisonnement, on voit que la probabilité de tirer blanche et noire, dans un ordre donné et sans remettre les boules après chaque tirage, est

$$\frac{a(a-1)\ldots(a-\mu+1)\,b(b-1)\ldots(b-\nu+1)}{(a+b)(a+b-1)\ldots(a+b-\mu-\nu+1)},$$

μ étant le nombre des blanches, ν celui des noires que l'on désire voir sortir.

Principe de la probabilité totale.

Nous appellerons *cause* d'un événement, dont l'arrivée n'est pas certaine, ce qui lui donne sa probabilité. Si, par exemple, une urne contient une boule blanche et une boule noire, la cause de l'extraction de la blanche ou de la noire sera la composition de l'urne en blanches et noires. Comme on le voit, le mot *cause* n'est pas détourné de son sens, il

reçoit une acception plus générale; la cause d'un événement certain sera ce qui lui donne la probabilité 1 ou sa certitude, ce sera bien sa CAUSE dans le sens propre du mot.

Nous dirons que des causes C_1, C_2, C_3,... *s'excluent mutuellement*, si l'événement attendu, arrivant sous l'influence de l'une d'elles C_i, on est sûr qu'aucune autre cause n'agira en même temps pour le produire.

THÉORÈME. — *Si un événement* E *peut être attribué à plusieurs causes* C_1, C_2,..., C_i,... *qui s'excluent mutuellement, sa probabilité* P *sera donnée par la formule*

$$P = p_1 q_1 + p_2 q_2 + \ldots + p_i q_i + \ldots,$$

où p_i *désigne, en général, la probabilité de* E *lorsque la cause* C_i *agit certainement et où* q_i *désigne la probabilite que cette cause est en jeu.*

Désignons par N le nombre total des cas également possibles qui peuvent se présenter quand on attend l'arrivée de E, par n_i le nombre des cas dans lesquels C_i agit et m_i le nombre des cas où, C_i agissant, E a lieu. Remarquons bien que les causes C_1, C_2,... s'excluent et, par conséquent, le nombre total de cas où E peut se présenter est

$$m_1 + m_2 + \ldots + m_i + \ldots;$$

il serait inférieur à $m_1 + \ldots + m_i + \ldots$, si les causes ne s'excluaient pas, parce que l'on compterait en trop les cas où E aurait lieu à la fois sous l'influence de deux ou plusieurs causes. Le nombre total des cas favorables à l'arrivée de E étant donc $m_1 + m_2 + \ldots + m_i + \ldots$, la probabilité cherchée P sera

$$P = \frac{m_1 + m_2 + \ldots + m_i + \ldots}{N},$$

ce que l'on peut écrire

$$P = \frac{m_1}{n_1}\frac{n_1}{N} + \frac{m_2}{n_2}\frac{n_2}{N} + \ldots + \frac{m_i}{n_i}\frac{n_i}{N} + \ldots;$$

mais, en général, $\frac{m_i}{n_i}$ est la probabilité p_i que, C_i agissant, E aura lieu, puisque c'est le rapport du nombre des cas favorables à l'arrivée de E au nombre total des cas qui peuvent se présenter, quand C_i agit; $\frac{n_i}{N}$ est, par définition, la probabilité q_i de C_i; donc on a bien

$$P = p_1 q_1 + p_2 q_2 + \ldots + p_i q_i + \ldots,$$

ainsi que nous l'avions annoncé.

On voit que la démonstration précédente manquerait de solidité si les causes C_1, C_2,... pouvaient agir simultanément; il faudra donc, dans les applications, vérifier, avec le plus grand soin, si les causes en jeu s'excluent réellement.

Par le seul fait de l'action de quelques causes ou même de toutes les causes, il peut se faire que l'événement ait forcément lieu; alors certaines quantités p_1, p_2,... sont égales à l'unité; en d'autres termes, les nombres de cas favorables à l'action des causes sont précisément les nombres de cas favorables à l'arrivée de l'événement. On énonce quelquefois le théorème, dans ce cas particulier, comme il suit :

THÉORÈME II. — *Lorsque les cas favorables à l'arrivée d'un événement peuvent se présenter de plusieurs manières, qui s'excluent mutuellement, la probabilité de cet événement est égale à la somme des probabilités que l'événement se présentera de chacune de ces manières.*

On peut immédiatement en conclure que, la probabilité d'un événement étant p, celle de l'événement contraire

étant q, on a

$$p + q = 1;$$

car le fait de l'arrivée ou de la non-arrivée d'un événement est un fait certain; sa probabilité *a priori* est 1. L'arrivée ou la non-arrivée sont deux façons dont ce fait peut se présenter : sa probabilité est donc aussi $p + q$.

Ces principes sont tellement importants, que nous croyons devoir en donner une nouvelle démonstration avant d'aller plus loin.

Le produit $p_i q_i$ est (en vertu du principe de la probabilité composée) la probabilité que la cause C_i agira pour produire l'événement E et que, cette cause ayant agi, l'événement E aura lieu; si l'on pose

$$p_i q_i = r_i,$$

r_i sera la probabilité que l'événement E aura lieu sous l'influence de la seule cause C_i. Or on peut admettre que la probabilité cherchée de E ne dépend que des quantités $r_1, r_2, \ldots, r_i, \ldots$; mais alors on peut supposer que $r_1, r_2, \ldots$ sont des fractions réduites au même dénominateur et poser

$$\frac{n_1}{D} = r_1, \quad \frac{n_2}{D} = r_2, \ldots, \quad \frac{n_i}{D} = r_i, \ldots;$$

la probabilité cherchée ne dépendant, par hypothèse, que de $r_1, r_2, \ldots$, on peut changer la nature du problème que l'on résout et supposer que, quand C_i agit, le nombre des cas favorables qui peuvent se présenter est n_i, le nombre total des cas étant D; la probabilité cherchée est donc

$$\frac{n_1 + n_2 + \ldots}{D} \quad \text{ou} \quad r_1 + r_2 + \ldots, \qquad \text{C. Q. F. D.}$$

Il est essentiel d'observer que, de quelque manière que

l'on présente la démonstration, on est forcé de faire cette hypothèse : que la solution ne dépeud que des quantités p et q, ou bien que l'on peut rendre tous les cas également possibles, en changeant la nature ou l'énoncé du problème; ce qui est au fond la même hypothèse.

PROBLÈME I. — *Une urne contient a boules blanches, b boules noires, c boules rouges. Quelle est la probabilité de ne pas tirer une boule blanche?*

On peut ne pas tirer une boule blanche de deux manières, soit en tirant rouge, soit en tirant noire; les probabilités respectives de ces deux événements sont $\frac{b}{a+b+c}$ et $\frac{c}{a+b+c}$; la probabilité cherchée est donc

$$\frac{b}{a+b+c}+\frac{c}{a+b+c},$$

et l'on vérifie que cette quantité est égale à

$$\frac{a}{a+b+c},$$

qui est la probabilité de tirer une blanche, c'est-à-dire la probabilité de ne tirer ni rouge, ni noire.

PROBLÈME II. — *On a n boules dans une urne. On demande la probabilité qu'en en prenant un certain nombre au hasard ce nombre sera pair ou impair.*

Les manières dont on peut prendre un nombre pair de boules sont les cas où l'on en prendra 2, 4, 6, 8,...; les probabilités de ces événements simples sont respectivement les quotients de C_n^2, C_n^4, C_n^6,..., par la somme de toutes les combinaisons que l'on peut faire avec n objets, somme que l'on sait être égale à $(1+1)^n$ ou 2^n; la probabi-

lité d'extraire un nombre pair de boules sera donc

$$\frac{C_n^2+C_n^4+C_n^6+\ldots}{2^n}=\frac{(1+1)^n+(1-1)^n-1}{2^{n+1}}=\frac{2^n-1}{2^{n+1}}=\frac{1}{2}-\frac{1}{2^{n+1}};$$

la probabilité d'extraire un nombre impair est de même

$$\frac{C_n^1+C_n^3+C_n^5+\ldots}{2^n}=\frac{(1+1)^n-(1-1)^n}{2^{n+1}}=\frac{2^n}{2^{n+1}}=\frac{1}{2}.$$

La probabilité d'extraire un nombre impair de boules est donc plus considérable que celle d'extraire un nombre pair, mais la différence tend vers zéro quand le nombre total des boules augmente indéfiniment.

Problème III. — *Trouver la probabilité d'amener au moins une fois as en jetant deux dés.*

On raisonnerait mal si l'on disait : la probabilité d'amener as avec le premier dé est $\frac{1}{6}$, avec le second $\frac{1}{6}$ aussi et, comme l'événement peut avoir lieu avec le premier ou le second dé, la probabilité cherchée $\frac{1}{6}+\frac{1}{6}$ ou $\frac{1}{3}$. En effet, l'événement attendu a pour causes les jets de chacun des dés, mais ces causes ne s'excluent pas mutuellement, en ce sens que, si as arrive avec le premier dé, il peut arriver avec le second.

Quoi qu'il en soit, nous allons cependant appliquer le principe de la probabilité totale, en observant que l'on peut amener as de trois manières : 1° avec le premier dé, sans l'amener avec le second; 2° avec le second, sans l'amener avec le premier; 3° avec les deux. Ces trois manières s'excluent et donnent à l'événement observé les probabilités

respectives

$$\frac{1}{6}\frac{5}{6}, \quad \frac{1}{6}\frac{5}{6}, \quad \frac{1}{6}\frac{1}{6}$$

(Il est facile de voir que $\frac{1}{6}\frac{5}{6}$ est la probabilité d'amener as avec le premier dé seul; en effet, en considérant cet événement comme composé du concours de deux autres, l'arrivée de as avec le premier dé, dont la probabilité est $\frac{1}{6}$, et l'arrivée de tout autre point que as avec le second, dont la probabilité est $\frac{5}{6}$, on a bien $\frac{1}{6}\frac{5}{6}$ pour la probabilité en question, etc...).

La probabilité demandée sera donc

$$\frac{1}{6}\frac{5}{6}+\frac{1}{6}\frac{5}{6}+\frac{1}{6}\frac{1}{6}=\frac{11}{36}.$$

Problème IV. — *On choisit une urne au hasard, dans un tas de n urnes; la première contient a_1 blanches, et b_1 noires; la seconde a_2 blanches, et b_2 noires, etc.... On demande la probabilité d'extraire une boule blanche de l'urne, tirée au hasard.*

Chacune des urnes en question peut être considérée comme une cause, ayant pour probabilité $\frac{1}{n}$; ces causes s'excluent évidemment, car une boule ne peut sortir à la fois de deux urnes; par suite, la probabilité P d'extraire une blanche sera

$$\frac{1}{n}\frac{a_1}{a_1+b_1}+\frac{1}{n}\frac{a_2}{a_2+b_2}+\ldots,$$

ou bien

$$P=\frac{1}{n}\left(\frac{a_1}{a_1+b_1}+\frac{a_2}{a_2+b_2}+\ldots+\frac{a_n}{a_n+b_n}\right).$$

Si l'on avait mis toutes les boules dans une même urne, la probabilité de la sortie d'une blanche eût été

$$\frac{a_1 + a_2 + \ldots + a_n}{a_1 + a_2 + \ldots + b_1 + b_2 + \ldots + b_n} = p.$$

Cherchons le maximum et le minimum de P, en supposant n, Σa et Σb constants. Égalons à zéro la différentielle totale de P, nous aurons

$$\sum \frac{a\,db - b\,da}{(a+b)^2} = 0.$$

On a d'ailleurs

$$\Sigma da = 0, \quad \Sigma db = 0,$$

et la méthode des multiplicateurs fournit les équations suivantes :

$$(a_i + b_i)^2\lambda - a_i = 0, \quad (a_i + b_i)^2\mu - b_i = 0,$$

où l'on doit faire $i = 1, 2, \ldots, n$; on en conclut

$$\frac{a_1}{a_1 + b_1} = \frac{a_2}{a_2 + b_2} = \ldots = \frac{a_n}{a_n + b_n} = \frac{\Sigma a}{\Sigma a + \Sigma b} = p,$$

$$\frac{b_1}{a_1 + b_1} = \frac{b_2}{a_2 + b_2} = \ldots = \frac{b_n}{a_n + b_n} = \frac{\Sigma b}{\Sigma a + \Sigma b} = 1 - p.$$

La première série d'égalités entraîne la seconde et l'on voit que la probabilité sera maxima ou minima, quand le rapport du nombre des boules blanches au nombre des boules noires sera le même dans chaque urne.

Le mode de raisonnement que nous venons d'employer semble absurde au premier abord, en ce sens que $a_1, a_2, \ldots$ ne varient pas d'une manière continue; mais, peu importe, P est une fonction continue de $a_1, a_2, \ldots$, qui atteint son

maximum ou son minimum pour certaines valeurs de ses variables; si l'on donne aux variables des valeurs discontinues, il y aura maximum ou minimum, quand les valeurs discontinues seront précisément égales à celles que l'on trouve, en admettant la continuité.

En tout cas, la probabilité maxima ou minima que l'on peut obtenir est la même que si toutes les boules étaient dans la même urne. Toutefois la condition imposée à a_1, $a_2, \ldots$, d'être positifs ou tout au moins nuls, introduit une nouvelle discontinuité, qui fera que d'autres maxima ou minima s'introduiront dans la question.

Problème V. — *Deux urnes contiennent : l'une a boules blanches et b boules noires, l'autre a' blanches et b' noires. On tire k boules de la première urne, pour les mettre dans la seconde; on demande la probabilité de tirer une boule blanche de la seconde après cette opération. On suppose $k < a$ et $< b$.*

Plusieurs cas ou causes peuvent donner lieu à l'extraction d'une boule blanche : 1° on a tiré k boules blanches de la première urne et zéro boules noires; 2° on a tiré $k-1$ blanches et 1 noire; 3° on a tiré $k-2$ blanches et 2 noires, etc.... Les probabilités respectives de ces hypothèses sont, comme nous l'avons vu (p. 39),

$$\frac{C_a^k}{C_{a+b}^k}, \quad \frac{C_a^{k-1}C_b^1}{C_{a+b}^k}, \quad \frac{C_a^{k-2}C_b^2}{C_{a+b}^k}, \ldots;$$

dans le premier cas, la probabilité de tirer une blanche (si l'on était certain que ce cas s'est présenté) serait $\frac{a'+k}{a'+b'+k}$; dans le second cas, la probabilité de tirer une blanche serait $\frac{a'+k-1}{a'+b'+k}$, et ainsi de suite. La probabilité

cherchée est donc

$$\frac{C_a^k}{C_{a+b}^k}\,\frac{a'+k}{a'+b'+k}+\frac{C_a^{k-1}C_b^1}{C_{a+b}^k}\,\frac{a'+k-1}{a'+b'+k}+\ldots,$$

ou bien

$$(1)\left\{\begin{array}{l}[C_a^k(a'+k)+C_a^{k-1}C_b^1(a'+k-1)+C_a^{k-2}C_b^2(a'+k-2)+\ldots]\\ \quad :C_{a+b}^k(a'+b'+k).\end{array}\right.$$

Si le nombre k était plus grand que l'un des nombres a et b ou même plus grand que chacun d'eux, la marche à suivre pour résoudre la question serait la même, seulement il faudrait faire attention que, si, par exemple, on avait $k > a$, on ne pourrait plus supposer que, dans le tirage de la première urne, on a extrait k blanches; la première hypothèse à faire serait l'extraction de a blanches et de $k-a$ noires; la seconde serait l'extraction de $a-1$ blanches et de $k-a+1$ noires et ainsi de suite, et la formule (1) subsisterait encore dans ce cas, pourvu que l'on convînt que le symbole C_a^k représente zéro, toutes les fois que k est supérieur à a.

Nous pouvons vérifier notre formule, en supposant $b'=0$, $a'=0$; la probabilité devient alors

$$[k C_a^k + (k-1)C_a^{k-1}C_b^1 + (k-2)C_a^{k-2}C_b^2 + \ldots] : C_{a+b}^k k.$$

Si l'on y fait $k = 1, 2, 3, \ldots$, on trouve toujours $\frac{a}{a+b}$, ce qui devait être *a priori*, puisque la question revient, dans ce cas, à demander la probabilité de tirer une blanche de la première urne.

Résolution des questions de probabilités par le théorème de Bayes.

Je rappelle, avant d'aller plus loin, que l'on dit que les

causes d'un événement s'*excluent mutuellement*, quand, l'une d'elles agissant certainement, les autres n'ont plus d'influence sur l'arrivée de l'événement. Je rappelle aussi que la probabilité d'une cause est la probabilité qu'elle est en jeu. Cela posé, nous allons passer à la démonstration d'un théorème fondamental dû au géomètre anglais Bayes.

Théorème I. — *Soient* $p_1, p_2, \ldots, p_i, \ldots$ *les probabilités que des causes* $C_1, C_2, \ldots, C_i, \ldots$, *s'excluant mutuellement, donnent respectivement à l'événement* E. *Soient* $q_1, q_2, \ldots, q_i, \ldots$ *les probabilités de ces causes. Supposons maintenant que l'événement* E *ait été observé dans une épreuve, la probabilité* ϖ_i *que l'arrivée de l'événement observé est due à la cause* C_i *est donnée par la formule*

$$\varpi_i = \frac{p_i q_i}{p_1 q_1 + p_2 q_2 + \ldots + p_i q_i + \ldots}.$$

En effet, soit H_i la probabilité que, avant toute épreuve, E ait lieu sous l'influence de C_i; en vertu du principe de la probabilité composée, pour obtenir H_i, il faudra multiplier la probabilité que C_i est en jeu ou q_i par la probabilité que, C_i agissant, E aura lieu ou p_i; donc

$$(1) \qquad H_i = p_i q_i;$$

mais, en vertu du même principe de la probabilité composée, H_i s'obtient en multipliant la probabilité P de l'arrivée de E, considérée *a priori*, et d'une manière absolue par la probabilité que, cet événement ayant eu lieu, il soit dû à l'influence de C_i, probabilité que nous avons désignée par ϖ_i. On a donc aussi

$$H_i = P\varpi_i,$$

d'où l'on tire, en comparant avec la formule (1),

$$(2) \qquad P\varpi_i = p_i q_i;$$

mais P est la probabilité déterminée *a priori* de E; cette probabilité est donnée par la formule

$$P = p_1 q_1 + p_2 q_2 + \ldots + p_i q_i + \ldots,$$

qui résulte du principe de la probabilité totale. De cette formule et de la formule (2) on tire

$$(3) \qquad \varpi_i = \frac{p_i q_i}{p_1 q_1 + p_2 q_2 + \ldots + p_i q_i + \ldots},$$

ce qu'il fallait démontrer.

Quand les causes ont les mêmes probabilités, la formule (3) devient

$$\varpi_i = \frac{p_i}{p_1 + p_2 + \ldots}.$$

Théorème II. — *Les mêmes choses que dans le théorème précédent étant posées, la probabilité pour que l'événement* E', *auquel les causes* $C_1, C_2, \ldots, C_i, \ldots$ *donnent les probabilités* $p'_1, p'_2, \ldots, p'_i, \ldots$, *ait lieu quand on a observé* E, *est donnée par la formule*

$$(4) \qquad p'_1 \varpi_1 + p'_2 \varpi_2 + \ldots + p'_i \varpi_i + \ldots$$

ou

$$(5) \qquad \frac{p_1 p'_1 q_1 + p_2 p'_2 q_2 + \ldots + p_i p'_i q_i + \ldots}{p_1 q_1 + p_2 q_2 + \ldots + p_i q_i + \ldots}.$$

En effet, la formule (4) est la formule même du principe de la probabilité totale appliquée à l'événement E', et la formule (5) s'en déduit en remplaçant $\varpi_1, \varpi_2, \ldots, \varpi_i, \ldots$ par leurs valeurs tirées de la formule (3).

Problème I. — *Une première urne contient* a_1 *boules blanches et* b_1 *noires, une seconde urne contient* a_2 *boules blanches et* b_2 *noires, etc.: on tire de ces urnes une boule*

qui se trouve être blanche, quelle est la probabilité qu'elle sort de la première urne?

Les causes sont ici les compositions des urnes; les probabilités que ces causes donnent à l'événement observé sont $\frac{a_1}{a_1+b_1}$, $\frac{a_2}{a_2+b_2}$, ... Soit n le nombre total des urnes, la probabilité que la main se portera sur chacune d'elles, avant d'observer l'événement est $\frac{1}{n}$, et, par suite, l'application du théorème de Bayes donne, pour la probabilité que la boule est sortie de la première urne,

$$\frac{1}{n}\frac{a_1}{a_1+b_1}:\sum\frac{1}{n}\frac{a}{a+b}, \quad \text{ou} \quad \frac{a_1}{a_1+b_1}:\sum\frac{a}{a+b}.$$

Si $a+b$ est le même pour chaque urne, on a simplement $\frac{a_1}{\Sigma a}$, et, si l'urne n° 1 contient le plus de blanches, c'est d'elle qu'il est le plus probable que sort la boule en question, ce qu'indiquaient *a priori* les règles du bon sens.

Problème II. — *Une urne contient des numéros marqués* 1, 2, 3, 4, ...; *en tirant au hasard, on extrait les numéros $a, b, c, \ldots, l$; on demande, d'après cela, le nombre probable des numéros contenus dans l'urne.*

Les numéros a, b, ..., l extraits constituent un événement dont les causes sont les hypothèses que l'on peut faire sur le nombre des numéros contenus dans l'urne; supposons ce nombre égal à x, la probabilité que ce nombre est x sera donnée par la formule

$$\frac{p_x}{\Sigma p_x},$$

p_x désignant la probabilité que le nombre x donne *a priori* à la sortie des numéros a, b, ..., l, que nous supposerons

rangés par ordre de grandeur et croissants; le signe Σ doit s'étendre de l à l'infini, car on ne peut pas supposer $x < l$.

Calculons donc p_x; le nombre des cas possibles est C_x^n, n désignant le nombre des numéros tirés, et le nombre des cas favorables est 1; on a donc, pour la probabilité de l'hypothèse que nous avons faite,

$$\frac{\frac{1}{C_x^n}}{\sum_{x=l}^{x=\infty} \frac{1}{C_x^n}},$$

ou bien

$$\frac{1}{\frac{x(x-1)\ldots(x-n+1)}{l(l-1)\ldots(l-n+1)} + \frac{x(x-1)\ldots(x-n+1)}{(l+1)l\ldots(l-n+2)} + \ldots};$$

le dénominateur est une série convergente dont on peut trouver la valeur en partant de la formule

$$\frac{z^{l-n}}{1-z} = z^{l-n} + z^{l-n+1} + z^{l-n+2} + \ldots.$$

En intégrant n fois de zéro à z, la dernière intégration se faisant en prenant $z = 1$, on a

$$\frac{1}{(l-n+1)\ldots l} + \frac{1}{(l-n+2)\ldots(l+1)} + \ldots = \int \frac{z^{l-n}}{1-z} dz^n,$$

ou bien

$$= \int_0^1 \frac{z^{l-n}(1-z)^{n-2}}{1.2\ldots(n-1)} dz.$$

On peut ainsi calculer, par les fonctions Γ ou par les séries à volonté, la probabilité de chacune des hypothèses faites sur le nombre des numéros; l'hypothèse qui a le plus de poids est, comme l'on voit, celle qui correspond à la

plus petite valeur possible de x, valeur qui est $x = l$; et, dans ce cas, la probabilité de l'hypothèse est l'inverse de la série

$$1 + \frac{l - n + 1}{l + 1} + \frac{(l - n + 1)(l - n + 2)}{(l + 1)(l + 2)} + \ldots;$$

le cas de $l = n$ doit être remarqué; la probabilité est alors l'inverse de la série

$$1 + \frac{1}{l + 1} + \frac{1.2}{(l + 1)(l + 2)} + \frac{1.2.3}{(l + 1)(l + 2)(l + 3)} + \ldots,$$

dont la valeur approche sans cesse de l'unité à mesure que l croît.

Examen des cas où l'emploi du calcul infinitésimal devient nécessaire.

Jusqu'ici nous ne nous sommes occupés que de l'estimation des chances d'événements qui ne pouvaient se présenter que d'un nombre fini de manières; il y a des questions dans lesquelles le nombre des cas possibles et favorables peut croître indéfiniment lorsque l'on cherche à les rendre tous *également possibles*. Je suppose, par exemple, qu'on lance un dard très-aigu au hasard dans une cible de surface S; sur cette cible on a tracé un contour fermé renfermant l'aire s. Si l'on demande la probabilité que le dard pénétrera à l'intérieur du contour s, pour résoudre la question, il faudra décomposer les aires S et s en portions rectangulaires égales et infiniment petites; chacune d'elles aura la même chance d'être atteinte par le dard, mais leurs nombres sont entre eux comme S et s; par suite, la probabilité cherchée est $\frac{s}{S}$.

Lorsqu'il devient ainsi nécessaire de subdiviser les cas

à l'infini pour leur donner des *facilités* égales, le Calcul infinitésimal, déguisé ou non, s'introduit forcément dans l'analyse des hasards; zéro et 1 ne sont plus alors des symboles de *certitude absolue*. Si, par exemple, on suppose dans la question précédente que le contour s se réduise à un point, il est clair que le rapport $\frac{s}{S}$ sera zéro : il y aura *certitude morale* de ne point atteindre le point s avec le dard, bien que cet événement ne soit pas mathématiquement impossible.

Problème I. — *On casse au hasard une tige très-mince en trois morceaux, quelle est la probabilité que ces trois morceaux pourront servir à construire un triangle?*

Il est essentiel de dire ce que nous entendons par une tige cassée au hasard *en trois morceaux : nous voulons dire par là que l étant la longueur de la tige et a, b, c étant les trois morceaux, pourvu que l'on ait $a+b+c=l$, chacune des combinaisons obtenues, en attribuant à a, b, c des valeurs quelconques, peut se présenter avec la même facilité.*

Partageons la tige en un très-grand nombre $2n$ de parties égales, et admettons d'abord que la tige ne puisse se casser qu'aux points de division. Soient x, y, z les nombres de divisions contenues dans les trois morceaux; pour que le triangle soit possible, il faudra que l'on ait

$$x < y + z, \quad y < x + z, \quad z < x + y,$$

et, en remplaçant z par $2n - x - y$,

$$x < n, \quad y < n, \quad x + y > n.$$

Ces formules seront satisfaites en prenant $x = 2$ avec $y = n - 1$; $x = 3$ avec $y = n - 1$, $n - 2$; $x = 4$ avec $y = n - 1$, $n - 2$, $n - 3$,...; et enfin $x = n - 1$ avec

$y = n-1, n-2, \ldots, 3, 2$; le nombre des cas favorables est donc

$$1+2+3+\ldots+(n-2) = \frac{(n-2)(n-1)}{2}.$$

Quant aux cas possibles, ils correspondent aux valeurs

$$x=1 \text{ avec } y=1, 2, 3, \ldots, 2n-2,$$
$$x=2 \text{ avec } y=1, 2, 3, \ldots, 2n-3, \ldots;$$

leur nombre est

$$2n-2+2n-3+2n-4+\ldots+1 = \frac{(2n-2)(2n-1)}{2}:$$

la probabilité cherchée est donc

$$(1) \qquad \frac{(n-2)(n-1)}{(2n-2)(2n-1)}.$$

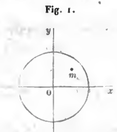
Si l'on suppose $n=\infty$, on aura la solution de la question en admettant que la tige puisse se briser en un point quelconque de sa longueur avec la même facilité; dans cette hypothèse, l'expression (1) se réduit à $\frac{1}{4}$.

Problème II. — *Une cible circulaire tourne uniformément autour d'un axe vertical, situé dans son plan et passant par son centre. On tire un coup de fusil dans une direction horizontale, et, en visant un point déterminé de la sphère engendrée par la cible dans son mouvement de rotation, en supposant l'arme juste, on demande la probabilité que l'on a d'atteindre la cible.*

Prenons l'axe de rotation pour axe des y et le plan vertical perpendiculaire à la trajectoire de la balle pour plan des xy; représentons sur ce plan le contour apparent de la sphère engendrée par la cible, ou, si l'on veut, la cible elle-même au moment où elle coïncide avec le plan en

question, l'axe des x passera par le centre de ce cercle. Soient m le point visé, x et y ses coordonnées. Nous supposerons la vitesse de la cible négligeable vis-à-vis de celle

Fig. 1.



du projectile, en sorte que la cible puisse être remplacée, sans erreur sensible, par sa projection sur le plan des xy. Soit θ l'angle de la cible avec le plan des xy au moment où la balle traverse le point m. La cible sera atteinte si sa projection contient le point m; elle ne le sera pas dans le cas contraire; en d'autres termes, la cible sera atteinte si le point m est intérieur à l'ellipse qui a pour équation

$$\frac{x^2}{\cos^2\theta} + y^2 = \mathrm{R}^2,$$

R désignant le rayon de la cible. Cette condition sera remplie si l'on a

$$\frac{x^2}{\cos^2\theta} + y^2 - \mathrm{R}^2 < 0,$$

d'où l'on tire

$$\cos^2\theta \geq \frac{x^2}{\mathrm{R}^2 - y^2}.$$

Le rapport du nombre des cas favorables au nombre des cas possibles sera alors le rapport de la valeur de l'arc déter-

miné par

$$\cos^2\theta = \frac{x^2}{R^2 - y^2}$$

à $\frac{\pi}{2}$, en ne prenant que la valeur de θ comprise dans le premier quadrant, et en ne considérant que le quart des nombres de cas favorables et possibles. La probabilité cherchée est donc

$$\frac{2}{\pi} \arccos \sqrt{\frac{x^2}{R^2 - y^2}} = f(x, y);$$

la probabilité $f(x, y)$ se change en certitude quand on a $x = 0$, et le lieu des points pour lesquels la probabilité reste la même est une ellipse, ce qui était facile à prévoir.

Problème III. — *Les mêmes choses étant posées que dans le problème précédent, supposons qu'au lieu de viser un point m déterminé on tire au hasard dans la cible; quelle sera la probabilité de l'atteindre?*

Pour résoudre cette question, décomposons le cercle tracé sur la figure en éléments rectangulaires infiniment petits de côtés dx et dy. La cible peut être atteinte pour une infinité de causes s'excluant; ces causes sont représentées par chacun des rectangles $dx\,dy$, leur probabilité commune est $\frac{dx\,dy}{\pi R^2}$; la cause $dx\,dy$ agissant, la probabilité de la rencontre a été calculée et trouvée égale à

$$\frac{2}{\pi} \arccos \sqrt{\frac{x^2}{R^2 - y^2}}:$$

donc la probabilité cherchée sera, en vertu du principe de la probabilité totale,

$$P = \frac{2}{\pi^2 R^2} \int\!\!\int \arccos \sqrt{\frac{x^2}{R^2 - y^2}}\, dx\, dy.$$

L'intégrale devant s'étendre à toute la surface du cercle de rayon R, pour effectuer l'intégration, nous conserverons la variable y, mais nous poserons

$$\sqrt{\frac{x^2}{R^2-y^2}} = \cos\theta,$$

nous aurons alors

$$P = -\frac{2}{\pi^2 R^2}\int\int \theta \sin\theta \sqrt{R^2-y^2}\,dy\,d\theta.$$

L'intégrale de $-\theta\sin\theta$ est $\theta\cos\theta - \sin\theta$, et, par suite,

$$\int \arccos\sqrt{\frac{x^2}{R^2-y^2}}\,dx$$
$$= \frac{x}{\sqrt{R^2-y^2}}\arccos\frac{x}{\sqrt{R^2-y^2}} - \sqrt{\frac{R^2-y^2-x^2}{R^2-y^2}};$$

prise entre les limites zéro et $\sqrt{R^2-y^2}$, cette quantité se réduit à l'unité; on a donc

$$P = \frac{4}{\pi^2 R^2}\int_{-R}^{+R} dy\sqrt{R^2-y^2} = \frac{4}{\pi^2 R^2}\frac{\pi R^2}{2} = \frac{2}{\pi}.$$

Supposons maintenant que l'adresse du tireur soit représentée par p; en d'autres termes, supposons que le tireur qui vise la cible ait la probabilité p de l'atteindre dans le cas où elle ne tourne pas; la probabilité qu'il aura de l'atteindre, quand elle tourne, sera celle d'un événement composé, et on l'obtiendra en multipliant la probabilité p d'atteindre la cible dans sa position directe par la probabilité $\frac{2}{\pi}$ de l'atteindre quand elle tourne et quand la balle frappe dans la sphère engendrée par la cible.

Ainsi, quand on tire dans une cible tournante placée à une distance assez grande pour qu'il y ait à peu près la même chance d'atteindre un quelconque de ses points lorsqu'elle ne tourne pas, la chance que l'on a de l'atteindre quand elle est en mouvement est à peu près les $\frac{2}{3}$ de celle que l'on avait de l'atteindre quand elle est au repos.

Problème IV. — *Deux individus* A, B *se donnent rendez-vous pour le même jour au même endroit; ils conviennent de s'attendre pendant une heure, sans préciser le moment où ils arriveront dans l'endroit en question; quelle est la probabilité qu'ils ont de se rencontrer?*

La rencontre peut avoir lieu de deux manières : ou A arrive le premier au rendez-vous, ou B y arrive le premier; chacun de ces événements a pour probabilité $\frac{1}{2}$; si l'on désigne par P la probabilité de la rencontre quand A arrive le premier, elle sera aussi P quand B arrivera le premier, et la probabilité de rencontre dans l'une ou l'autre hypothèse sera

$$\frac{1}{2}(P + P),$$

c'est-à-dire encore P. Nous pouvons donc supposer que A arrive le premier sans altérer la probabilité cherchée. Comptons le temps à partir de l'origine de la journée du rendez-vous et désignons-le par t; soit T la durée de la journée et τ celle de l'heure en fonction d'une unité de temps que nous laisserons indéterminée. A peut arriver à chacun des instants de la journée, et la probabilité qu'il arrive à une époque comprise entre t et $t + dt$ est $\frac{dt}{T}$; considérons l'arrivée de A comme une cause, la probabilité

de la rencontre, quand cette cause est un jeu, sera égale au rapport du nombre $\frac{\tau}{dt}$ des cas où B arrivera pendant le séjour de A au nombre total $\frac{T-t}{dt}$ des cas qui pourront se présenter, nombre qui est mesuré par l'intervalle compris entre les epoques t et T, puisque A est arrivé nécessairement avant B. La probabilité de la rencontre, en vertu du principe de la probabilité totale, sera la somme des produits

$$\frac{dt}{T} \times \frac{\tau}{T-t} \quad \text{ou} \quad \int \frac{\tau\, dt}{T(T-t)}.$$

Toutefois, il faut observer que si A arrive aux époques comprises entre $T-\tau$ et T, la rencontre est certaine : il faudra donc prendre l'intégrale précédente entre les limites zéro et $T-\tau$, et y ajouter l'intégrale de $\frac{dt}{T} \times 1$ ou de $\frac{dt}{T}$ prise entre les limites $T-\tau$ et T; car, entre ces limites, la probabilité de la rencontre n'est plus $\frac{\tau}{T-\tau}$, mais bien l'unité. On a ainsi

$$P = \int_0^{T-\tau} \frac{\tau\, dt}{T(T-t)} + \int_{T-\tau}^{T} \frac{dt}{T},$$

ou bien

$$(1) \qquad P = -\frac{\tau}{T} \log \frac{\tau}{T} + \frac{\tau}{T};$$

la probabilité cherchée ne dépend donc que du rapport $\frac{\tau}{T}$. Prenons la journée de douze heures, nous aurons

$$\frac{\tau}{T} = \frac{1}{12} \quad \text{et} \quad P = \frac{1}{12} + \frac{1}{12} \log 12,$$

En effectuant les calculs, on a à peu près

$$P = 0,2904.$$

On peut contrôler l'exactitude de la formule (1) en prenant $\tau = T$; dans ce cas on a un rendez-vous ordinaire entre gens exacts. La rencontre est certaine, on doit donc avoir $P = 1$: effectivement, si l'on fait $T = \tau$ dans (1), on trouve bien $P = 1$. De même, si l'on y fait $\tau = 0$, on trouve $P = 0$, et, en effet, on conçoit que, si les séjours de A et de B sont infiniment courts, ils risqueront fort de ne pas se rencontrer.

On a ici un exemple de probabilité nulle qui n'indique pas une impossibilité absolue, mais une impossibilité morale. Effectivement, les individus A et B peuvent passer au même instant au rendez-vous, mais cet événement en lui-même n'a aucune chance en sa faveur.

La considération de l'infini s'introduit quelquefois d'une manière différente dans les questions de probabilités. Ainsi il peut se faire que l'application du principe de la probabilité totale ou de la probabilité composée conduisent à des séries ou à des produits infinis. Nous allons en donner quelques exemples.

Problème V. — *Un individu tire dans une urne deux fois de suite, cette urne contient une blanche et une noire : s'il tire deux fois blanche, il gagne 1 franc, s'il ne tire pas deux fois blanche, on le laisse tirer encore deux fois de suite dans une urne qui contient deux noires et une blanche, et ainsi de suite, en augmentant toujours de 1 le nombre des boules noires. Quelle est la probabilité que cet individu a de gagner son franc?*

Il est plus commode, pour résoudre la question, de chercher la probabilité que l'individu a de ne pas gagner le

franc, parce que c'est chercher la probabilité que le tirage se continuera indéfiniment; nous avons alors affaire à un événement composé des suivants : 1° que dans la première urne l'individu ne tirera pas 2 fois blanche de suite : la probabilité de tirer blanche 2 fois de suite est $\frac{1}{2}\times\frac{1}{2}$ ou $\frac{1}{4}$, et la probabilité de ne pas tirer 2 fois de suite blanche est $\left(1-\frac{1}{4}\right)$; 2° que dans la deuxième urne l'individu en question ne tirera pas 2 fois de suite blanche : la probabilité de cet événement simple est $1-\frac{1}{9}$, et ainsi de suite; la probabilité cherchée est donc

$$\left(1-\frac{1}{4}\right)\left(1-\frac{1}{9}\right)\left(1-\frac{1}{16}\right)\cdots\left(1-\frac{1}{n^2}\right)\cdots;$$

or on a

$$\sin x = x\left(1-\frac{x^2}{\pi^2}\right)\left(1-\frac{x^2}{4\pi^2}\right)\left(1-\frac{x^2}{9\pi^2}\right)\cdots\left(1-\frac{x^2}{n^2\pi^2}\right)\cdots,$$

d'où l'on tire

$$\frac{\sin x}{x\left(1-\frac{x^2}{\pi^2}\right)} = \left(1-\frac{x^2}{4\pi^2}\right)\left(1-\frac{x^2}{9\pi^2}\right)\cdots,$$

et, en faisant $x=\pi$,

$$\frac{1}{2\pi}\lim\frac{\sin x}{\left(1-\frac{x}{\pi}\right)} = \left(1-\frac{1}{4}\right)\left(1-\frac{1}{9}\right)\cdots,$$

ou

$$\frac{1}{2} = \left(1-\frac{1}{4}\right)\left(1-\frac{1}{9}\right)\cdots.$$

Ainsi la probabilité que l'individu a de ne pas gagner

le franc est $\frac{1}{2}$, la probabilité qu'il a de le gagner est donc $1-\frac{1}{2}$ ou également $\frac{1}{2}$.

Problème VI. — *On dit que des joueurs jouent une poule quand deux d'entre eux font une partie, le perdant se trouvant remplacé par un autre joueur, jusqu'à ce que le même joueur ait gagné deux fois de suite; ceci posé, trois joueurs* A, B, C *jouent une poule; on demande la probabilité que chacun a de la gagner, en supposant que* A *et* B *commencent d'abord.*

A peut gagner la poule :

1° En gagnant la première partie; 2° en la perdant.

1° Supposons que A gagne la première partie, ce dont la probabilité est $\frac{1}{2}$, il pourra gagner la poule dans l'une des hypothèses comprises dans le tableau suivant :

	1re hypothèse.	2e hypothèse.	3e hypothèse.
1re partie....	A gagne.	A gagne.	A gagne...
2e »	A gagne.	B gagne.	B gagne...
3e »	....	C gagne.	C gagne...
4e »	....	A gagne.	A gagne...
5e »	....	A gagne.	B gagne...
6e »	....	....	C gagne...
7e »	....	....	A gagne...
8e »	....	....	A gagne...

Les probabilités respectives de ces hypothèses sont $\frac{1}{2}$, $\frac{1}{2^4}$, $\frac{1}{2^7}$, $\frac{1}{2^{10}}$, ...; donc la probabilité que A gagnera la poule,

après avoir gagné la première partie, est

$$P = \frac{1}{2}\left(\frac{1}{2} + \frac{1}{2^4} + \frac{1}{2^7} + \frac{1}{2^{10}} + \ldots\right) = \frac{1}{4}\,\frac{1}{1 - \frac{1}{2^3}} = \frac{2}{7}.$$

2° Supposons que A perde la première partie, ce dont la probabilité est $\frac{1}{2}$; en dressant un tableau comme le précédent, on verrait que la probabilité qu'il a de gagner la poule est

$$P' = \frac{1}{2}\left(\frac{1}{2^3} + \frac{1}{2^6} + \frac{1}{2^9} + \ldots\right) = \frac{1}{2^4}\,\frac{1}{1 - \frac{1}{2^3}} = \frac{1}{14}.$$

La probabilité totale que A a de gagner est donc $P + P'$ ou $\frac{5}{14}$; c'est aussi la probabilité qu'a B de gagner la poule; la probabilité qu'a C de gagner est

$$1 - \frac{5}{14} - \frac{5}{14} \quad \text{ou} \quad \frac{2}{7}.$$

Disons enfin que l'infini peut encore s'introduire dans certaines questions, lorsqu'y laissant des données indéterminées on fait tendre ces données vers certaines limites.

Supposons, par exemple, que l'on demande la probabilité de tirer m fois de suite une boule noire d'une urne qui contient m boules, dont une blanche et $m - 1$ noires.

La probabilité d'extraire une boule noire est $\frac{m-1}{m}$ ou $1 - \frac{1}{m}$, la probabilité de l'extraire m fois de suite est

$$\left(1 - \frac{1}{m}\right)^m.$$

Si l'on suppose que l'on prenne le nombre m de plus en plus grand, cette quantité tend, comme l'on sait, vers la limite $\frac{1}{e}$, de telle sorte que, si m est très-grand, on peut estimer sans grande erreur la probabilité cherchée égale à $\frac{1}{e}$; il n'est pas d'ailleurs sans intérêt de faire quelquefois de ces hypothèses qui ne se réalisent pas, mais qui servent à fixer des limites entre lesquelles varient les quantités que l'on cherche; ainsi, dans l'exercice précédent, on peut voir que, $\left(1-\frac{1}{m}\right)^m$ allant en croissant avec m, la probabilité cherchée restera toujours inférieure à $\frac{1}{e}$, c'est-à-dire environ à $\frac{10}{27}$.

Résolution des questions de probabilité par la méthode des fonctions génératrices.

Nous avons vu que l'on pouvait traiter les questions d'analyse combinatoire en considérant les résultats du calcul comme les termes de certains développements; cette méthode est très-fréquemment appliquée dans le calcul des probabilités. Nous allons en donner quelques exemples.

Problème I. — *Étant donnés n dés présentant chacun f faces, numérotées 1, 2, 3,..., f, calculer la probabilité d'amener un nombre de points α avec ces dés.*

On connaît la toupie vulgairement appelée *tonton;* sa forme générale est celle d'une surface prismatique, terminée par une tige et une pointe; sur chaque face de cette toupie est marqué un numéro; lorsque le mouvement imprimé à ce petit jouet s'éteint, il tombe sur une de ses

faces; un dé de f faces sera, si l'on veut, un tonton dont la surface prismatique présentera f pans égaux.

Ceci posé, désignons par $a, b, c, \ldots, l$ les numéros inscrits sur les faces sur lesquelles les dés tombent; les cas favorables à l'arrivée de l'événement attendu sont les manières dont on peut choisir les nombres $a, b, \ldots, l$, de telle sorte qu'ils satisfassent à l'égalité

$$a + b + \ldots + l = \alpha;$$

or, si l'on considère la fonction

$$(x + x^2 + x^3 + \ldots + x^f)^n = \mathrm{X},$$

le coefficient de x^α dans son développement sera le nombre de manières dont on pourra choisir les exposants de x compris entre 1 et f, de telle sorte que la somme de n d'entre eux fasse α. Ainsi le coefficient de x^α sera le numérateur de la probabilité cherchée; quant au dénominateur, il sera évidemment égal à la somme de tous les coefficients des diverses puissances de x, somme que l'on peut obtenir en faisant $x = 1$ dans la fonction X, et que l'on trouve ainsi égale à f^n.

Or on a

$$\mathrm{X} = \left(x \frac{1 - x^f}{1 - x}\right)^n,$$

ou bien

$$\mathrm{X} = x^n (1 - x^f)^n (1 - x)^{-n};$$

les binômes qui figurent dans cette expression se développent par les formules

$$(1 - x^f)^n = 1 - \frac{n}{1} x^f + \frac{n(n-1)}{1.2} x^{2f} - \ldots,$$

$$(1 - x)^{-n} = 1 + \frac{n}{1} x + \frac{n(n+1)}{1.2} x^2 + \ldots,$$

d'où l'on conclut facilement le coefficient de x^α dans X. En désignant par N_α ce coefficient, la probabilité cherchée devient

$$\frac{N_\alpha}{f^n}.$$

Supposons que l'on demande la probabilité d'amener 7 avec deux dés ordinaires, cette probabilité sera le coefficient de x^7 dans

$$\frac{1}{36}x^2(1-x^6)^2(1-x)^{-2},$$

ou de x^5 dans

$$\frac{1}{36}(1-x^6)^2(1-x)^{-2}$$

$$=\frac{1}{36}(1-2x^6+x^{12})(1+2x+3x^2+4x^3+5x^4\ldots);$$

le coefficient de x^5 est

$$\frac{6}{36}=\frac{1}{6}.$$

Le point 7 est le point qui a le plus de chance de se présenter; en effet, les probabilités des points 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 sont respectivement entre elles comme les nombres 1, 2, 3, 4, 5, 6, 5, 4, 3, 2, 1; il y a plus, on voit: 1° qu'il y a autant de chances d'amener un point supérieur à 7 que d'en amener un inférieur; 2° que deux points dont la somme fait 14 ont des chances égales de se présenter; cette conclusion peut être généralisée et étendue à un nombre quelconque de dés présentant le même nombre de faces. La démonstration de ce théorème est facile à établir, en convertissant en intégrale définie le terme général du développement de X, conformément à la méthode donnée

(p. 5), ou bien encore en observant que les termes également distants des extrêmes dans X doivent avoir des coefficients égaux; on pourrait presque en conclure que le terme du milieu est un maximum, et s'en servir pour calculer le point le plus probable.

Problème II. — *Une urne contient des boules numérotées* 1, 2, 3,..., n; *on en tire* α *d'un seul coup, et l'on demande la probabilité pour que la somme des numéros tirés soit égale à* s.

La probabilité cherchée est celle d'un événement dans lequel les chances favorables peuvent se présenter de plusieurs manières qui s'excluent, à savoir de toutes les manières dont on peut grouper α des nombres 1, 2, 3,..., n, de manière que leur somme fasse s; or le coefficient de $t^\alpha x^s$ dans

$$f(x, t) = (1 + tx)(1 + tx^2)(1 + tx^3)\ldots(1 + tx^n)$$

est aussi le nombre des manières dont on peut prendre parmi les exposants 1, 2,..., n, de x, α nombres dont la somme soit s. Le nombre des cas favorables à l'événement dont nous cherchons la probabilité est donc le coefficient de $t^\alpha x^s$ dans $f(x, t)$, le nombre des cas possibles est C_n^α, d'où l'on conclura facilement la probabilité quand on connaîtra le coefficient de $t^\alpha x^s$. Euler a donné le coefficient de t^α sous forme explicite, lorsque $n = \infty$, dans son *Introductio in Analysin;* mais je ne crois pas que l'on connaisse l'expression de ce coefficient dans le cas où n est fini, en fonction de n.

Nous pourrions pousser plus loin l'étude de cette question; nous ne le ferons pas, parce que notre but dans ce Chapitre a moins été de résoudre des questions difficiles que de montrer comment on devait appliquer et com-

prendre les principes. Dans les Chapitres suivants, nous appliquerons successivement le calcul des probabilités à l'étude des phéuomènes et de leurs causes, aux jeux de toute espèce, et en particulier aux opérations des compagnies d'assurances, qui sont de véritables jeux.

Méthodes indirectes.

La probabilité d'un événement peut souvent se présenter comme l'inconnue d'une équation à résoudre. Nous allons montrer le genre de questions qui conduisent à des solutions de ce genre.

Problème I. — *Trois joueurs font une poule* (*voir* la définition de ce mot, p. 71), A *joue d'abord contre* B *et perd ; on demande quelle est alors pour chaque joueur la probabilité de gagner la poule.*

Soient respectivement x, y, z les probabilités relatives à A, B, C, on aura d'abord

$$x + y + z = 1,$$

car l'un des joueurs gagnera. Or B d'abord peut gagner : 1° en gagnant la première partie : la probabilité de cette hypothèse est $\frac{1}{2}$; 2° en perdant cette première partie, mais en gagnant deux parties de suite aux coups suivants, ce dont la probabilité est x, car alors il se trouve dans le même cas que A au début. On a donc, en observant que la probabilité de la seconde hypothèse est $\frac{1}{2}$,

$$y = \frac{1}{2} + \frac{1}{2} x.$$

Quant à C, il ne peut gagner la poule qu'en gagnant la

première partie qu'il joue ; la probabilité de cette hypothèse est $\frac{1}{2}$, mais alors il se trouve dans le même cas que B au début, et la probabilité de gagner la poule en gagnant la première partie est $\frac{1}{2}y$; on a donc

$$z = \frac{1}{2}y;$$

on tire des trois équations trouvées

$$x = \frac{1}{7}, \quad y = \frac{4}{7}, \quad z = \frac{2}{7}.$$

Si l'on demandait la probabilité que chaque joueur a de gagner au début, on raisonnerait comme il suit : A jouant contre B peut gagner ou perdre la première partie ; s'il la perd, la probabilité qu'il a de gagner est $\frac{1}{7}$; s'il la gagne, la probabilité qu'il a de gagner est la probabilité que B avait de gagner tout à l'heure ou $\frac{4}{7}$; donc la probabilité que A a de gagner la poule est, avant de commencer la première partie,

$$\frac{1}{2}\frac{1}{7} + \frac{1}{2}\frac{4}{7},$$

c'est-à-dire $\frac{5}{14}$; c'est aussi la probabilité relative à B. Quant à la probabilité relative à C, elle est $1 - 2\frac{5}{14}$ ou $\frac{2}{7}$, ainsi que nous l'avions trouvé plus haut.

Si au lieu de trois joueurs on considérait un plus grand nombre, la marche à suivre pour résoudre la question serait la même.

PROBLÈME II. — *Deux joueurs* A, B *possèdent l'un a francs l'autre b francs; ils jouent* 1 *franc la partie; on demande la probabilité que* A *ruinera* B *ou que* B *ruinera* A.

Soit $f(x)$ la probabilité que A sera ruiné lorsqu'il possède x francs. A peut être ruiné soit en gagnant, soit en perdant la partie qui va se jouer. S'il gagne, il possédera $x+1$ francs et la probabilité qu'il sera ruiné sera $f(x+1)$; s'il perd, il possédera $x-1$ francs et sa probabilité d'être ruiné sera $f(x-1)$; mais la ruine de A étant un événement qui peut avoir lieu sous l'influence de deux causes distinctes dont la probabilité est $\frac{1}{2}$ pour chacune d'elles, à savoir qu'il gagnera ou qu'il perdra, on aura, en vertu du principe de la probabilité totale,

$$f(x)=\frac{1}{2}f(x+1)+\frac{1}{2}f(x-1),$$

d'où l'on tire

$$f(x+1)-2f(x)+f(x-1)=0.$$

La différence seconde de $f(x)$ est donc nulle; donc sa différence première est constante : je dis constante et non périodique, parce que nous n'avons évidemment à considérer que des valeurs entières de x; mais la différence première de $f(x)$ étant constante et égale à c, par exemple, on a

$$f(x)=f(x-1)+c,$$
$$f(x-1)=f(x-2)+c,$$
$$\dots\dots\dots\dots\dots\dots,$$
$$f(1)=f(0)+c.$$

Or $f(0)$ est égal à 1; on a donc, en ajoutant,

$$f(x)=1+cx.$$

En faisant $x = a$ et $x = b$ dans cette formule, on a

$$f(a) = 1 + ca, \quad f(b) = 1 + cb;$$

or $f(a)$ et $f(b)$ ont pour somme l'unité, car l'un des joueurs A et B sera ruiné; en ajoutant alors les deux formules précédentes, on trouve

$$1 = 2 + c(a + b);$$

cette relation fait connaître c et l'on a

$$c = -\frac{1}{a+b};$$

il en résulte

$$f(a) = 1 - \frac{a}{a+b}, \quad f(b) = 1 - \frac{b}{a+b},$$

ou, si l'on veut,

$$f(a) = \frac{b}{a+b}, \quad f(b) = \frac{a}{a+b}.$$

On voit ainsi, ce que le simple bon sens indique d'ailleurs, que le joueur le plus riche est celui qui a le moins de chances d'être ruiné; il y a plus, si l'on fait $a = \infty$, on trouve $f(b) = 1$, et le joueur qui joue contre un individu infiniment plus riche que lui est inévitablement ruiné; ce joueur infiniment riche existe, c'est le public. Le joueur de profession joue contre le public, et la discussion précédente met en évidence le désavantage de sa position; toujours d'accord avec les principes de la morale, c'est-à-dire du bon sens, le calcul des probabilités fournit souvent des préceptes et de sages conseils dans notre conduite.

Un grand nombre de questions, relatives au calcul des probabilités, conduisent ainsi à des équations aux différences finies; malheureusement les méthodes que l'on a

pour intégrer ces équations sont fondées sur des théories que nous ne pouvons pas exposer ici, parce qu'elles ne comportent pas une rigueur suffisante, et c'est peut-être ce qui a empêché les analystes modernes de se livrer à l'étude du Calcul des probabilités. On ne voit pas, en effet, que cette science ait fait des progrès sensibles depuis Laplace et Poisson, du moins en dehors de ses applications aux méthodes d'observation.

Dans le Traité de Laplace, on peut lire un certain nombre de solutions de problèmes qui conduisent à des équations aux différences finies partielles; ces questions sont résolues par la méthode des fonctions génératrices. On pourrait souvent représenter directement la solution au moyen d'intégrales définies, mais les résultats ainsi obtenus ne satisfont point assez l'esprit pour que nous en présentions ici des exemples.

Problème III. — A *joue avec* B; *à chaque partie, les probabilités qu'ils ont respectivement de gagner sont* p *et* q, *en sorte que* $p+q=1$; *ils possèdent en entrant au jeu* a *francs et* b *francs; on demande la probabilité que* A *ruinera* B *avant le coup de rang* n.

Soit $f(x,y)$ la probabilité que A ruinera B quand ce dernier possède encore x francs et quand on a encore y coups à jouer. Deux causes peuvent amener la ruine de B : 1° le gain de la partie qui va se jouer par A; 2° le gain de la même partie par B. Les probabilités de ces causes sont p et q respectivement; mais la première cause donne à la ruine de B la probabilité $f(x-1, y-1)$, la seconde lui donne la probabilité $f(x+1, y-1)$; on a donc, en vertu du principe de la probabilité totale,

$$(1)\qquad f(x,y)=pf(x-1,y-1)+qf(x+1,y-1).$$

On intégrera cette équation aux différences en posant

$$f(x, y) = ca^x b^y,$$

c étant arbitraire, et a et b étant simplement assujettis à vérifier l'équation

$$(1\ bis) \qquad 1 = pa^{-1}b^{-1} + qab^{-1},$$

ou

$$b = qa + pa^{-1};$$

on pourra donc prendre

$$f(x, y) = ca^x(qa + pa^{-1})^y,$$

ou, plus généralement,

$$f(x, y) = \Sigma ca^x(qa + pa^{-1})^y,$$

le signe Σ s'étendant à une suite indéfinie de valeurs arbitraires de c et de a. Développant alors $(qa + pa^{-1})^y$ par la formule du binôme de Pascal, et posant

$$\Sigma ca^x = \varphi(x),$$

on aura

$$(2) \quad \begin{cases} f(x, y) = q^y\varphi(x + y) + \frac{y}{1} q^{y-1}p\varphi(x + y - 2) \\ \qquad + \frac{y(y-1)}{1.2} q^{y-2}p^2\varphi(x + y - 4) + \dots \end{cases}$$

Maintenant observons que $f(x, y)$ est nul pour $y = 0$, si B possède quelque chose, c'est-à-dire tant que x est plus grand que zéro. Au contraire, $f(x, y)$ est égal à l'unité, si

x est nul quand $y = 0, 1, 2, 3, \ldots$; en ayant égard à ces deux remarques, l'équation (2) fournit les relations

$$(3) \qquad 0 = \varphi(x) \text{ pour } x = 1, 2, 3, \ldots;$$

$$(4) \qquad 1 = q^y \varphi(y) + \frac{y}{1} q^{y-1} p \varphi(y-2) + \ldots + p^y \varphi(-y);$$

la relation (4) pourra donc s'écrire, en vertu de la relation (3),

$$1 = p^y \varphi(-y) + \frac{y}{1} p^{y-1} q \varphi(-y+2) \ldots,$$

et dans cette formule on ne doit prendre que les termes où l'argument sous le signe φ est nul ou négatif; on a alors, pour $y = 0, 1, 2, 3, \ldots$,

$$1 = \varphi(0), \qquad \text{ou} \qquad \varphi(0) = 1;$$

$$1 = p\,\varphi(-1), \qquad \varphi(-1) = \frac{1}{p};$$

$$1 = p^2 \varphi(-2) + 2pq\varphi(0), \qquad \varphi(-2) = \frac{1-2pq}{p^2};$$

$$1 = p^3 \varphi(-3) + 3p^2 q \varphi(-1), \qquad \varphi(-3) = \frac{1-3pq}{p^3};$$

$$\ldots\ldots\ldots\ldots, \qquad \ldots\ldots\ldots\ldots$$

On pourra ainsi calculer de proche en proche les valeurs de $\varphi(0)$, $\varphi(-1)$, $\varphi(-2)$, ... La formule (2) fera alors connaître $f(x, y)$. Mais on peut donner une autre solution de la question; et, en effet, reprenant la formule (1 *bis*) qui lie entre elles les constantes a et b, on en tire

$$a = \frac{p}{b} + \frac{qa^2}{b};$$

6.

faisant alors usage de la formule de Lagrange pour calculer a^x, on a

$$a^x = \left(\frac{p}{b}\right)^x + \frac{x}{1}\frac{p^{x+1}q}{b^{x+2}} + \frac{x(x+3)}{1.2}\frac{p^{x+2}q^2}{b^{x+4}} + \frac{x(x+3)(x+5)}{1.2.3}\frac{p^{x+3}q^3}{b^{x+6}} + \ldots.$$

La solution de l'équation (1) étant de la forme

$$ca^x b^y, \quad \text{ou, plus généralement,} \quad \Sigma ca^x b^y,$$

on pourra écrire

$$f(x, y) = \sum c\left(\frac{p}{b}\right)^x b^y + \sum c\,\frac{x}{1}\,\frac{p^{x+1}q}{b^{x+2}}\,b^y + \ldots,$$

ou bien, en posant Σcb^u égal à $\varphi(u)$,

$$(5)\quad \begin{cases} f(x, y) = p^x\Big[\varphi(y-x) + \dfrac{x}{1}pq\varphi(y-x-2) \\ \qquad + \dfrac{x(x+3)}{1.2}p^2q^2\varphi(y-x-4) + \ldots\Big]. \end{cases}$$

Or nous avons déjà vu que $f(x, 0) = 0$ pour $x > 0$, et $f(0, y) = 1$ pour $y = 0, 1, 2, 3, \ldots$; l'équation (5) donne alors

$$0 = \varphi(-x) + \frac{x}{1}pq\varphi(-x-2) + \ldots,$$
$$1 = \varphi(y).$$

La première de ces formules sera satisfaite en posant $\varphi(x) = 0$ pour toutes les valeurs négatives de x; la seconde exige, au contraire, que $\varphi(0)$, $\varphi(1)$, $\varphi(2)$,... soient égaux

à l'unité; on a donc, au lieu de l'équation (5),

$$f(x, y) = p^x \left[1 + xpq + \frac{x(x+3)}{1.2} p^2 q^2 + \frac{x(x+3)(x+5)}{1.2.3} p^3 q^3 + \ldots \right],$$

le développement devant être poussé jusqu'à ce que $y - x$ devienne égal ou supérieur au double de l'exposant de q.

Maintenant, pour résoudre la question proposée, il suffit de faire $x = b$ et $y = n$, et l'on a, pour la probabilité cherchée,

$$f(b, n) = p^b \left[1 + bpq + \frac{b(b+3)}{1.2} p^2 q^2 + \ldots \right].$$

La question que nous venons de traiter montre bien où se trouvent généralement concentrées les principales difficultés du Calcul des probabilités; et il est évident que les progrès de ce Calcul dépendent en grande partie de ceux de la théorie des équations aux différences finies.

Observation générale relative à l'application des principes exposés dans ce Chapitre.

L'application des principes exposés ci-dessus est assez délicate, et les personnes qui abordent le Calcul des probabilités feront bien de vérifier que les principes s'appliquent réellement aux questions qu'elles cherchent à résoudre, en recommençant sur ces questions particulières les démonstrations des principes fondamentaux.

Voici, par exemple, un problème auquel une application trop précipitée du principe de la probabilité composée conduirait à un résultat inexact.

Problème. — *Un étang contient un nombre de poissons*

d'égale grosseur et de même espèce représenté par n; on demande la probabilité de pêcher, en deux coups de filet, d'abord a puis b poissons.

L'application faite à la légère du principe de la probabilité composée conduirait à dire : « La probabilité de pêcher d'abord a poissons est

$$C_n^a : \sum_{m=0}^{m=n} C_n^m \quad \text{ou} \quad C_n^a : 2^n;$$

la probabilité d'en pêcher b quand on en a déjà pêché a est

$$C_{n-a}^b : \sum_{m=0}^{m=n-a} C_{n-a}^m \quad \text{ou} \quad C_{n-a}^b : 2^{n-a},$$

et, par suite, la probabilité cherchée est

$$\frac{C_n^a . C_{n-a}^b}{2^{2n-a}} \text{ »}.$$

Mais il est facile de voir que le principe de la probabilité composée ne s'applique pas à ce problème, parce qu'à chaque cas possible de la première épreuve ne correspond pas toujours le même nombre de cas possibles de la seconde épreuve. Si, par exemple, on pêche k poissons au premier coup de filet, on en pourra pêcher soit 0, soit 1, soit 2, soit 3,..., soit $n-k$ au second coup, ce qui constitue en définitive

$$1 + C_{n-k}^1 + C_{n-k}^2 + \ldots + C_{n-k}^{n-k} = 2^{n-k} \text{ cas possibles;}$$

ce nombre de cas est donc essentiellement variable, suivant le cas qui se présente au premier coup de filet. Du reste, en traitant le problème directement, on reconnaît que

$C_n^a . C_{n-a}^b$ est le nombre des cas favorables; celui des cas possibles s'obtient en observant que l'on peut pêcher o poissons au premier coup de filet, et au second coup il pourra se présenter 2^n cas différents; on peut pêcher 1 poisson au premier coup de C_n^1 manières, auxquelles correspondent 2^{n-1} cas au second coup, etc. En général, on peut pêcher k poissons au premier coup, et cela de C_n^k manières, à chacune desquelles correspondent, comme nous l'avons vu tout à l'heure, 2^{n-k} cas au second coup de filet; en résumé, le nombre des cas possibles est

$$2^n + C_n^1 2^{n-1} + C_n^2 2^{n-2} + \ldots + C_n^k 2^{n-k} + \ldots + C_n^n = 3^n;$$

on a donc, pour la probabilité cherchée,

$$\frac{C_n^a . C_{n-a}^b}{3^n}.$$

CHAPITRE III.

ÉTUDE DES PHÉNOMÈNES QUE L'ON OBSERVE DANS LA RÉPÉTITION DES MÊMES ÉPREUVES.

Probabilité des événements composés de la répétition des mêmes événements simples.

Problème I. — *Soit* p *la probabilité d'un événement* E, q *la probabilité de l'événement contraire* F; *calculer la probabilité pour que dans un nombre donné* s *d'épreuves l'événement* E *ait lieu* α *fois et* F, $s-\alpha$ *fois.*

Supposons d'abord que l'ordre dans lequel E et F doivent avoir lieu soit donné, et cherchons la probabilité pour que les événements E et F se présentent : E d'abord a fois de suite; puis F, b fois de suite; puis E, c fois de suite, etc. Cette probabilité est celle d'un événement composé qui dépend du concours de plusieurs autres E, E,...; F, F,...; E, E,..., dont les arrivées ne se gênent en aucune façon; elle est donc égale au produit des probabilités des événements simples en question ou $p^a q^b p^c \ldots$, c'est-à-dire en intervertissant l'ordre des facteurs $p^\alpha q^{s-\alpha}$, ou plus simplement en remplaçant $s-\alpha$ par β, $p^\alpha q^\beta$.

Supposons maintenant que l'on fasse abstraction de l'ordre dans lequel les événements E et F se présentent dans les diverses épreuves : la probabilité cherchée est alors celle d'un événement qui peut se présenter de plusieurs manières s'excluant les unes les autres; ce nombre de manières est

celui dont on peut grouper α lettres identiques à E et β lettres identiques à F (β désignant toujours $s-\alpha$).

Il résulte de là, et du principe de la probabilité totale (p. 47), que la probabilité cherchée est la somme des probabilités $p^\alpha q^\beta$ obtenues en assignant un ordre déterminé à l'avance aux événements E, F, et cela de toutes les manières possibles. La solution cherchée s'obtiendra donc en répétant $p^\alpha q^\beta$ autant de fois qu'il y a d'unités dans le nombre de manières d'arranger deux lettres E, F; en en prenant α identiques à E et β identiques à F, nous avons trouvé ce nombre (Chap. I, p. 14 et 15) égal à $\frac{s!}{\alpha!\,\beta!}$ ou à C_s^α, et par suite la probabilité pour que l'événement E ait lieu α fois et F, β fois, abstraction faite de l'ordre où ces événements se présentent, est

$$C_s^\alpha p^\alpha q^\beta = \frac{s!}{\alpha!\,\beta!} p^\alpha q^\beta.$$

Remarquons dès à présent que cette probabilité est le $\alpha+1^{ième}$ terme du développement de $(p+q)^s$, ordonné suivant les puissances croissantes de p. Si α et β sont de très-grands nombres, on évaluera le résultat en remplaçant $\alpha!$ par sa valeur approchée $\alpha^\alpha e^{-\alpha}\sqrt{2\pi\alpha}$, déduite de la formule de Stirling (p. 8).

PROBLÈME II. — *Dans une série de s épreuves, l'un des événements* E, F, G, H,... *doit nécessairement se présenter à chaque épreuve; les probabilités respectives de ces événements sont p, q, r,..., en sorte que*

$$p+q+r+\ldots=1.$$

On demande la probabilité pour que E *se présente α fois,* F, *β fois,* G, *γ fois, etc.*

En reprenant les raisonnements faits plus haut, on voit

que la probabilité pour que ces événements E, F, G,... aient lieu dans un ordre fixé à l'avance est

$$p^{\alpha} q^{\beta} r^{\gamma} \ldots .$$

Si l'on fait abstraction de l'ordre, l'événement attendu pourra se présenter d'autant de manières que l'on peut grouper α lettres identiques à E avec β lettres identiques à F, avec γ lettres identiques à G, etc., c'est-à-dire de $\frac{s!}{\alpha!\,\beta!\,\gamma!}$ manières (p. 15), et, par suite, la probabilité cherchée sera

$$\frac{s!}{\alpha!\,\beta!\,\gamma!\ldots} p^{\alpha} q^{\beta} r^{\gamma} \ldots .$$

C'est précisément le terme en $p^{\alpha} q^{\beta} r^{\gamma}\ldots$ dans $(p+q+r+\ldots)^{s}$.

PROBLÈME III. — *Quelle est la probabilité pour que dans une série de s épreuves l'événement* E, *dont la probabilité est p, ait lieu un nombre de fois compris entre* $\alpha - l$ *et* $\alpha + l$?

La probabilité demandée est la somme des probabilités que cet événement aura lieu $\alpha - l$ fois, $\alpha - l + 1$ fois, ..., α fois, $\alpha + 1$ fois,..., $\alpha + l$ fois, c'est-à-dire, d'après ce que nous venons de voir,

$$\begin{aligned} & C_s^{\alpha-l} p^{\alpha-l} q^{\beta+l} + C_s^{\alpha-l+1} p^{\alpha-l+1} q^{\beta+l-1} + \ldots \\ & + C_s^{\alpha} p^{\alpha} q^{\beta} + \ldots + C_s^{\alpha+l} p^{\alpha+l} q^{\beta-l}, \end{aligned}$$

q désignant toujours la quantité $1 - p$. La probabilité demandée est donc, comme l'on voit, la somme des termes du développement de $(p+q)^{s}$ compris entre le terme en $p^{\alpha+l}$ et le terme en $p^{\alpha-l}$ inclusivement.

Corollaire. — La probabilité pour que l'événement E ait lieu au moins n fois, c'est-à-dire pour qu'il ait lieu

n fois, ou $n+1$ fois, ou $n+2$ fois, ... est

$$C_s^n p^n q^{s-n} + C_s^{n+1} p^{n+1} q^{s-n-1} + \ldots;$$

c'est la somme des $s-n+1$ premiers termes du développement de $(p+q)^s$ ordonné par rapport aux puissances *croissantes de* q. On verrait de même que la probabilité pour que E ait lieu au plus n fois est la somme des $n+1$ premiers termes de $(p+q)^s$, ordonné par rapport aux puissances croissantes de p.

Problème IV. — *Soit toujours* p *la probabilité de l'événement* E, *et* $q=1-p$; *on demande quel est, dans une série de* s *épreuves, le nombre de fois que l'événement* E *a le plus de chance de se présenter.*

Pour résoudre cette question, il suffit de se rappeler que les termes du développement de $(p+q)^s$ sont respectivement les probabilités que l'événement E aura lieu 0, 1, 2,..., s fois dans s épreuves. Pour trouver le nombre de fois que cet événement a le plus de chance de se présenter, il suffit donc de chercher le plus grand terme du binôme $(p+q)^s$, soit

$$(1) \qquad \frac{s!}{\alpha!\,\beta!}\,p^\alpha q^\beta;$$

ce plus grand terme, le précédent et le suivant sont respectivement

$$(2) \qquad \frac{s!}{(\alpha-1)!\,(\beta+1)!}\,p^{\alpha-1}q^{\beta+1}, \quad \frac{s!}{(\alpha+1)!\,(\beta-1)!}\,p^{\alpha+1}q^{\beta-1}.$$

En écrivant que l'expression (1) est plus grande que chacune des expressions (2), on trouve

$$\frac{p^{\alpha-1}q^{\beta+1}}{(\alpha-1)!\,(\beta+1)!} < \frac{p^\alpha q^\beta}{\alpha!\,\beta!} > \frac{p^{\alpha+1}q^{\beta-1}}{(\alpha+1)!\,(\beta-1)!},$$

ou bien

$$\frac{p}{\alpha} > \frac{q}{\beta+1}, \quad \frac{p}{\alpha+1} < \frac{q}{\beta};$$

on en conclut

$$\frac{\beta+1}{\alpha} > \frac{q}{p} > \frac{\beta}{\alpha+1};$$

en ajoutant l'unité de part et d'autre, et en observant que $p+q=1$, on a

$$\frac{s+1}{\alpha} > \frac{1}{p} > \frac{s+1}{\alpha+1}.$$

On voit ainsi que

$$(s+1)p > \alpha, \quad \text{mais} \quad (s+1)p < \alpha+1,$$

et, par suite, α est le plus grand entier contenu dans $p(s+1)$, de même β est le plus grand entier contenu dans $q(s+1)$, et l'on a à peu près, quand s est très-grand,

$$\alpha = ps, \quad \beta = qs;$$

par suite, α et β sont sensiblement proportionnels à p et q. On serait déjà tenté de conclure de là que, si l'on fait un très-grand nombre d'épreuves, l'événement E et l'événement contraire se présenteront des nombres de fois proportionnels à leurs probabilités respectives; mais il faut observer que le plus grand terme du binôme, qui correspond à l'événement composé le plus probable, n'a qu'une valeur très-petite, et pour nous en assurer, il suffit de calculer ce terme au moyen des formules approchées (p. 8):

$$s! = s^s e^{-s} \sqrt{2\pi s};$$

$$\alpha! = \alpha^\alpha e^{-\alpha} \sqrt{2\pi\alpha}, \quad \beta! = \beta^\beta e^{-\beta} \sqrt{2\pi\beta};$$

$$p = \frac{\alpha}{s}, \quad q = \frac{\beta}{s}.$$

On trouve alors, pour la valeur du terme en question,

$$\sqrt{\frac{s}{2\pi\alpha\beta}}\,\frac{s^s}{\alpha^\alpha\beta^\beta}\left(\frac{\alpha}{s}\right)^\alpha\left(\frac{\beta}{s}\right)^\beta,$$

c'est-à-dire, à fort peu de chose près,

$$\sqrt{\frac{s}{2\pi\alpha\beta}} \quad \text{ou} \quad \sqrt{\frac{1}{2\pi pq}}\,\frac{1}{\sqrt{s}}.$$

Lorsque s est grand, on voit que l'événement composé le plus probable l'est fort peu en lui-même, ainsi que nous l'avions annoncé.

Le simple bon sens aurait pu faire prévoir une partie des résultats qui précèdent; mais ce qu'il était impuissant à révéler, c'est la manière dont les probabilités des événements se développent à mesure que l'on s'éloigne de l'événement le plus probable.

Bernoulli (Jacques) paraît être le premier qui ait attiré l'attention des géomètres sur cette question; il a énoncé, en effet, un théorème remarquable dans son *Ars conjectandi*, théorème qui a été beaucoup perfectionné depuis, surtout par Laplace et Condorcet. Voici l'énoncé du théorème de Bernoulli :

Théorème de Bernoulli.

THÉORÈME DE BERNOULLI. — *Soit p la probabilité d'un événement simple* E, α *le nombre de fois qu'il se présente dans une série de s épreuves; soit* P *la probabilité que la différence entre p et $\frac{\alpha}{s}$ sera inférieure en valeur absolue à ε. On peut toujours prendre s assez grand pour que* P *diffère de l'unité d'aussi peu que l'on voudra.*

Nous allons d'abord démontrer le théorème de Bernoulli ; dans le paragraphe suivant nous indiquerons la manière de calculer la quantité P avec telle exactitude que l'on voudra.

Nous avons vu tout à l'heure que si l'on posait $q=1-p$, $\beta=s-\alpha$,

$$(1)\quad \Pi = C_s^{\alpha-l} p^{\alpha-l} q^{\beta+l} + \ldots = \sum_{i=-l}^{i=+l} \frac{s!}{(\alpha-i)!\,(\beta+i)!} p^{\alpha-i} q^{\beta+i}$$

représentait la probabilité que l'événement E aura lieu un nombre de fois compris entre $\alpha-l$ et $\alpha+l$. Supposons que α représente le nombre de fois que E a le plus de chance d'arriver; comme l'on sait, α est le plus grand entier compris dans $(s+1)p$, en sorte que l'on peut poser

$$\alpha = (s+1)p - \varepsilon,$$

ε étant compris entre 0 et 1, ou

$$\alpha = sp + \omega,$$

ω étant moindre que 1 en valeur absolue; de même

$$\beta = sq + \omega',$$

et, comme on doit avoir $\alpha+\beta=s$, on en conclut $\omega'=-\omega$; ainsi

$$(2)\quad p = \frac{\alpha}{s} - \frac{\omega}{s}, \quad q = \frac{\beta}{s} + \frac{\omega}{s}.$$

La formule de Stirling donne

$$s! = s^s e^{-s} \sqrt{2\pi s}\, e^{\varpi(s)},$$

$$(\alpha-i)! = (\alpha-i)^{\alpha-i} e^{-\alpha+i} \sqrt{2\pi(\alpha-i)}\, e^{\varpi(\alpha-i)},$$

$$(\beta+i)! = (\beta+i)^{\beta+i} e^{-\beta-i} \sqrt{2\pi(\beta+i)}\, e^{\varpi(\beta+i)};$$

portant ces valeurs et celles de p et de q tirées de la formule (2) dans la formule (1), on a

$$(3)\quad \left\{ \begin{aligned} \Pi &= \sum_{-l}^{+l} \frac{s^s}{(\alpha-i)^{\alpha-i}(\beta+i)^{\beta+i}} \sqrt{\frac{s}{2\pi(\alpha-i)(\beta+i)}} \\ &\quad\times e^{\Omega}\left(\frac{\alpha}{s}-\frac{\omega}{s}\right)^{\alpha-i}\left(\frac{\beta}{s}+\frac{\omega}{s}\right)^{\beta+i}, \end{aligned} \right.$$

et Ω désigne, pour abréger, $\varpi(s)-\varpi(\alpha-i)-\varpi(\beta+i)$. Supposons $s=\infty$, $\alpha=\infty$, $\beta=\infty$, $l=\infty$; mais supposons le rapport $\frac{l}{\alpha}$ infiniment petit, en sorte que $\frac{i}{\alpha}$ et $\frac{i}{\beta}$ soient infiniment petits; Ω sera infiniment petit, et l'on pourra écrire

$$\begin{aligned} \Pi &= \sum_{-l}^{+l}\left(\frac{\alpha}{\alpha-i}\right)^{\alpha-i}\left(\frac{\beta}{\beta+i}\right)^{\beta+i}\sqrt{\frac{s}{2\pi\alpha\beta}}\sqrt{\frac{\alpha\beta}{(\alpha-i)(\beta+i)}} \\ &\quad\times e^{\Omega}\left(1-\frac{\omega}{\alpha}\right)^{\alpha-i}\left(1+\frac{\omega}{\beta}\right)^{\beta+i}; \end{aligned}$$

or on a

$$\begin{aligned} \log\left(\frac{\alpha}{\alpha-i}\right)^{\alpha-i} &= -(\alpha-i)\log\left(1-\frac{i}{\alpha}\right) \\ &= (\alpha-i)\left(\frac{i}{\alpha}+\frac{i^2}{2\alpha^2}+\frac{i^3}{3\alpha^3}+\ldots\right) \\ &= i-\frac{i^2}{2\alpha}-\frac{i^3}{6\alpha^2}-\ldots, \end{aligned}$$

de même

$$\log\left(\frac{\beta}{\beta+i}\right)^{\beta+i} = -i-\frac{i^2}{2\beta}+\frac{i^3}{6\beta^2}+\ldots,$$

d'où l'on conclut aisément

$$\log\left(\frac{\alpha}{\alpha-i}\right)^{\alpha-i}\left(\frac{\beta}{\beta+i}\right)^{\beta+i} = -\frac{i^2}{2\alpha}-\frac{i^2}{2\beta}+\varepsilon.$$

Si l'on suppose $\frac{i^2}{2\alpha}$ et $\frac{i^2}{2\beta}$ finis, ε sera infiniment petit, et l'on aura, en observant que $\alpha + \beta = s$,

$$(4) \qquad \left(\frac{\alpha}{\alpha - i}\right)^{\alpha - i}\left(\frac{\beta}{\beta + i}\right)^{\beta + i} = e^{-\frac{i^2 s}{2\alpha\beta} + \varepsilon}.$$

Le rapport $\sqrt{\frac{\alpha\beta}{(\alpha - i)(\beta + i)}}$ a pour limite l'unité; il en est de même de $\left(1 - \frac{\omega}{\alpha}\right)^{\alpha - i}$ et de $\left(1 + \frac{\omega}{\beta}\right)^{\beta + i}$; leur produit sera de la forme $e^{\varepsilon'}$, ε' ayant pour limite zéro. En désignant alors par μ une quantité moindre que le maximum de $\Omega + \varepsilon + \varepsilon'$, l'équation (3) deviendra, en ayant égard à l'équation (4),

$$\Pi = e^{\mu} \sum_{-l}^{+l} \sqrt{\frac{s}{2\pi\alpha\beta}}\, e^{-\frac{i^2 s}{2\alpha\beta}};$$

à la limite, e^{μ} est égal à 1, et l'on a

$$\lim \Pi = \lim \sum_{-l}^{+l} \sqrt{\frac{s}{2\pi\alpha\beta}}\, e^{-\frac{i^2 s}{2\alpha\beta}}.$$

On peut maintenant poser

$$\sqrt{\frac{s}{2\alpha\beta}} = d\gamma \quad \text{et} \quad i\sqrt{\frac{s}{2\alpha\beta}} = \gamma,$$

et l'on trouve

$$\lim \Pi = \frac{1}{\sqrt{\pi}} \int_{-g}^{+g} e^{-\gamma^2} d\gamma,$$

g désignant pour abréger

$$l\sqrt{\frac{s}{2\alpha\beta}} \quad \text{ou} \quad \frac{l}{\sqrt{2pqs}}.$$

et l'on peut encore écrire

$$\lim \Pi = \frac{2}{\sqrt{\pi}} \int_0^{\frac{l}{\sqrt{2pqs}}} e^{-\gamma^2} d\gamma.$$

Or Π représente la probabilité que E se présentera un nombre de fois f compris entre $\alpha - l$ et $\alpha + l$, ou, ce qui revient au même, que $\frac{f}{s}$ sera compris entre $\frac{\alpha}{s} - \frac{l}{s}$ et $\frac{\alpha}{s} + \frac{l}{s}$, c'est-à-dire entre $p - \frac{l}{s}$ et $p + \frac{l}{s}$; or on peut prendre $\frac{l}{\sqrt{s}} = \infty$ et $\frac{l}{s} = 0$, et cela d'une infinité de manières; alors on aura

$$\lim \Pi = \frac{2}{\sqrt{\pi}} \int_0^{\infty} e^{-\gamma^2} d\gamma = 1,$$

et l'on aura la probabilité 1 ou la certitude que le rapport $\frac{f}{s}$ du nombre de fois que E se présentera au nombre total des épreuves s ne diffère pas de p. Cela démontre le théorème de Bernoulli. Il y a plus, l'intégrale

$$\frac{2}{\sqrt{\pi}} \int_0^{\frac{l}{\sqrt{2pqs}}} e^{-\gamma^2} d\gamma$$

est une valeur approchée de la probabilité que E arrivera un nombre de fois compris entre $\alpha - l$ et $\alpha + l$.

La loi des grands nombres, ou la généralisation du théorème de Bernoulli.

Lemme. — Soit $f(\theta)$ une fonction réelle ou imaginaire

de la variable réelle θ, on a évidemment

$$f(\theta)-f(o)=\int_o^\theta f'(\theta-z)dz$$

ou

$$f(\theta)=f(o)+\int_o^\theta f'(\theta-z)dz,$$

et, en intégrant par parties, on trouve

$$f(\theta)=f(o)+\theta f'(o)+\frac{\theta^2}{2}f''(o)+\frac{1}{2}\int_o^\theta f'''(\theta-z)z^2dz.$$

Si dans l'intégrale du second nombre de cette formule on remplace $f'''(\theta-z)$ par son module maximum μ entre les limites o et θ de z, on a

$$\text{mod.}\,\frac{1}{2}\int_o^\theta f'''(\theta-z)z^2dz<\frac{1}{2}\int_o^\theta \mu z^2dz \quad \text{ou} \quad <\mu\frac{\theta^3}{6},$$

et μ représentera aussi le module maximum de $f'''(\theta)$ entre les limites o et θ; en désignant donc par ε une imaginaire dont le module est plus petit que l'unité, on pourra écrire

$$(1) \qquad f(\theta)=f(o)+\theta f'(o)+\frac{\theta^2}{2}f''(o)+\varepsilon\mu\frac{\theta^3}{6}.$$

Cette formule va bientôt nous être utile.

Loi des grands nombres. — Poisson a donné le nom de *loi des grands nombres* à la loi suivant laquelle un événement, dont la probabilité n'est pas nécessairement constante à chaque épreuve, se répète dans un grand nombre d'épreuves. L'expression rigoureuse de cette loi n'est pas connue. Les efforts de Poisson ne sont cependant pas restés infructueux, et il est parvenu à généraliser le théorème de

Jacques Bernoulli, en suivant d'assez près une analyse que l'on peut lire dans le *Traité des probabilités* de Laplace, et qui est relative aux bénéfices des Compagnies d'assurances.

Désignons par $p_1, p_2, \ldots, p_s$ (*) les probabilités d'un événement E aux épreuves nos 1, 2,..., s, et par $q_1, q_2, \ldots, q_s$ les probabilités de l'événement contraire F, en sorte que

$$p_1 + q_1 = 1, \ldots \quad p_s + q_s = 1.$$

Le principe de la probabilité composée nous apprend que la probabilité de l'arrivée de E aux épreuves nos $1, 2, 3, \ldots, \alpha$, de l'événement F aux épreuves nos $\alpha + 1$, $\alpha + 2, \ldots, \beta$, de l'événement E aux épreuves nos $\beta + 1$, $\beta + 2, \ldots, \gamma, \ldots$, est $p_1 p_2 \ldots p_\alpha q_{\alpha+1} q_{\alpha+2} \ldots q_\beta p_{\beta+1} p_{\beta+2} \ldots p_\gamma \ldots$. Si l'on n'assigne pas aux événements E et F un ordre déterminé à l'avance, la probabilité P que sur s épreuves l'événement E aura lieu α fois, et l'événement contraire $\beta = s - \alpha$ fois, en vertu du principe de la probabilité totale, sera égale à la somme des produits obtenus en prenant α facteurs p et β facteurs q de toutes les manières possibles. Cette probabilité sera donc le coefficient de $e^{\alpha\theta\sqrt{-1}}$ dans le développement de

$$\prod (q + pe^{\theta\sqrt{-1}}) = (q_1 + p_1 e^{\theta\sqrt{-1}})(q_2 + p_2 e^{\theta\sqrt{-1}}) \ldots (q_s + p_s e^{\theta\sqrt{-1}}),$$

ordonné suivant les puissances de $e^{\theta\sqrt{-1}}$. Ce coefficient est

(*) Si l'on considère une urne contenant des boules blanches et noires, l'extraction d'une boule blanche sera un événement dont la probabilité variera à chaque épreuve, si l'on modifie aussi à chaque épreuve le nombre des boules contenues dans l'urne. Dans une Compagnie d'assurances, les sinistres sont des événements dont la probabilité est variable avec la nature de l'assurance relative à chaque sinistre; c'est surtout à cause des applications que nous ferons au sujet de ces Compagnies que nous exposons la loi des grands nombres.

égal, d'après ce que nous avons vu p. 5, à

$$(1)\qquad P = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \prod\left(q + pe^{\theta\sqrt{-1}}\right)e^{-x\theta\sqrt{-1}}d\theta.$$

Pour évaluer cette intégrale, commençons par développer le logarithme de $\prod\left(q + pe^{\theta\sqrt{-1}}\right)$ suivant les puissances de θ; nous aurons d'abord, en posant

$$\log\left(q + pe^{\theta\sqrt{-1}}\right) = f(\theta),$$

en vertu de la formule démontrée au lemme précédent,

$$(2)\qquad \log\left(q + pe^{\theta\sqrt{-1}}\right) = f(0) + \theta f'(0) + \frac{\theta^2}{2}f''(0) + \varepsilon\frac{\mu\theta^3}{6};$$

or on a

$$(3)\qquad f'(\theta) = \frac{pe^{\theta\sqrt{-1}}\sqrt{-1}}{q + pe^{\theta\sqrt{-1}}} = \frac{p\sqrt{-1}}{qe^{-\theta\sqrt{-1}} + p},$$

$$(4)\qquad f''(\theta) = \frac{-pqe^{-\theta\sqrt{-1}}}{\left(qe^{-\theta\sqrt{-1}} + p\right)^2} = \frac{-pq}{\left(qe^{-\frac{\theta}{2}\sqrt{-1}} + pe^{\frac{\theta}{2}\sqrt{-1}}\right)^2},$$

$$(5)\qquad \begin{cases} f'''(\theta) = \sqrt{-1}\,\dfrac{pq\left(pe^{\frac{\theta}{2}\sqrt{-1}} - qe^{-\frac{\theta}{2}\sqrt{-1}}\right)}{\left(qe^{-\frac{\theta}{2}\sqrt{-1}} + pe^{\frac{\theta}{2}\sqrt{-1}}\right)^3} \\[2ex] \qquad = \sqrt{-1}\,\dfrac{pq\left(\overline{p-q}\cos\frac{\theta}{2} + \overline{p+q}\sin\frac{\theta}{2}\sqrt{-1}\right)}{\left(\overline{p+q}\cos\frac{\theta}{2} + \overline{p-q}\sin\frac{\theta}{2}\sqrt{-1}\right)^3}. \end{cases}$$

Les formules (3) et (4) donnent pour $\theta = 0$, en observant que $p + q = 1$,

$$f'(0) = p\sqrt{-1};\quad f''(0) = -pq;$$

quant à la formule (5), elle donnera

$$\text{mod.} f'''(\theta) = pq\sqrt{\frac{p^2+q^2-2pq\cos\theta}{(p^2+q^2+2pq\cos\theta)^3}}.$$

Si au numérateur et au dénominateur nous supprimons $2pq\cos\theta$, nous aurons évidemment, en supposant $\theta < \frac{\pi}{2}$,

$$\text{mod.} f'''(\theta) < \frac{pq}{(p^2+q^2)}.$$

Le maximum de $\frac{pq}{p^2+q^2}$ est $\frac{1}{2}$; on a donc

$$\text{mod.} f'''(\theta) < 1,$$

et la formule (2) devient, en appelant μ une quantité dont le module est moindre que 1,

$$\log\left(q+pe^{\theta\sqrt{-1}}\right) = p\theta\sqrt{-1} - pq\frac{\theta^2}{2} + \mu\frac{\theta^3}{6}.$$

Nous poserons

$$\sum p - \alpha = g, \quad \sum \frac{pq}{2} = k^2;$$

nous aurons ainsi

$$(6) \qquad \prod\left(q+pe^{\theta\sqrt{-1}}\right)e^{-\alpha\theta\sqrt{-1}} = e^{g\theta\sqrt{-1}-k^2\theta^2+\mu\frac{\theta^3}{6}}.$$

Cela posé, revenons à la formule (1); remplaçons dans l'intégrale qu'elle contient les limites $-\pi$ et $+\pi$ par $-h$ et $+h$, en désignant par h une quantité dont nous fixerons plus loin la valeur. Nous pourrons, en vertu de la formule (6), écrire

$$(7) \qquad \text{P} = \frac{1}{2\pi}\int_{-h}^{+h} e^{g\theta\sqrt{-1}-k^2\theta^2+\mu\frac{\theta^3}{6}}d\theta + \Omega,$$

Ω désignant une quantité évidemment inférieure à celle que

l'on obtient en remplaçant dans le second membre de la formule (1) : 1° la quantité placée sous le signe $\int$ par son module; 2° en faisant abstraction de la portion de l'intégrale relative aux éléments compris entre $-h$ et $+h$. On aura donc

$$\Omega < \frac{1}{\pi}\int_h^\pi \prod \text{mod.}\left(q + pe^{\theta\sqrt{-1}}\right)d\theta$$

ou bien

$$\Omega < \frac{1}{\pi}\int_h^\pi \prod \left(1 - 4pq\sin^2\frac{\theta}{2}\right)^{\frac{1}{2}} d\theta.$$

Si l'on observe que $\sin x$ est toujours plus grand que $\frac{x}{2}$ entre les limites 0 et $\frac{\pi}{2}$, on aura

$$\Omega < \frac{1}{\pi}\int_h^\pi \prod \left(1 - \frac{pq}{4}\theta^2\right)^{\frac{1}{2}} d\theta,$$

et, *a fortiori*, entre les limites h et π (*),

$$\Omega < \frac{1}{\pi}\int_h^\pi \prod e^{-\frac{pq}{8}\theta^2} d\theta,$$

ou bien

$$\Omega < \frac{1}{\pi}\int_h^\infty e^{-k^2\frac{\theta^2}{4}} d\theta \quad \text{ou} \quad < \frac{2}{\pi k}\int_{\frac{hk}{2}}^\infty e^{-t^2}dt.$$

(*) La différence entre $1 - \frac{pq}{4}\theta^2$ et $e^{-\frac{pq}{4}\theta^2}$ est $\left(\frac{pq\theta^2}{4}\right)^2\frac{1}{1.2} - \ldots$ Cette série est à termes positifs et négatifs, et décroissants quand $\theta = \pi$. En effet, pq est moindre que $\frac{1}{4}$, et $\frac{pq\theta^2}{4}$ moindre que $\frac{\pi^2}{16}$, quantité inférieure à l'unité; les propriétés bien connues de la série exponentielle nous apprendront donc que les termes vont en décroissant, et par suite l'erreur, quand on s'arrête à un terme, est moindre que ce terme.

Il suffit que h ne soit pas très-petit pour que l'intégrale qui figure dans cette formule soit excessivement faible, car k est de l'ordre de $\sqrt{s}$; en effet k^2 désigne la somme de s quantités telles que pq, et que nous ne supposerons pas indéfiniment décroissantes, quand s croît. L'intégrale $\int_\gamma^\infty e^{-t^2}dt$ étant désignée par $\mathrm{T}(\gamma)$, on aura donc

$$\Omega < \frac{2}{\pi k}\mathrm{T}\left(\frac{hk}{2}\right);$$

en appelant alors λ une quantité comprise entre 0 et 1, on pourra écrire, au lieu de la formule (7),

$$\mathrm{P} = \frac{1}{2\pi}\int_{-h}^{+h} e^{g\theta\sqrt{-1}-k^2\theta^2+\mu\frac{\theta^3}{6}}d\theta + \frac{2\lambda}{\pi k}\mathrm{T}\left(\frac{hk}{2}\right).$$

Si h est, au plus, égal à l'unité, on pourra écrire, en désignant toujours par μ' une quantité de module inférieur à l'unité,

$$e^{\mu\frac{\theta^3}{6}} = 1 + \frac{\mu'\theta^3}{6}$$

et l'on aura

$$(8)\qquad \mathrm{P} = \frac{1}{2\pi}\int_{-h}^{+h} e^{g\theta\sqrt{-1}-k^2\theta^2}d\theta + \omega + \frac{2\lambda}{\pi k}\mathrm{T}\left(\frac{hk}{2}\right),$$

d'où

$$\omega = \frac{1}{2\pi}\int_{-h}^{+h} e^{g\theta\sqrt{-1}-k^2\theta^2} \times \frac{\mu'\theta^3}{6}d\theta;$$

remplaçant les variables par leurs modules ou par des quantités plus grandes,

$$\omega < \frac{1}{\pi}\int_{0}^{+h} e^{-k^2\theta^2} \times \frac{\theta^3}{6}d\theta,$$

c'est-à-dire

$$\omega < \frac{1}{\pi k}\int_0^{hk} e^{-t^2} \times \frac{t^3}{6k^3}\,dt,$$

ou

$$< \frac{1}{6k^4\pi}\int_0^{\infty} e^{-t^2} t^3\,dt, \quad \text{ou} \quad < \frac{1}{3\pi k^4}.$$

Si l'on désigne encore par λ' une quantité moindre que 1, on aura

$$(9) \qquad \omega = \frac{\lambda'}{3k^4\pi}.$$

Enfin si l'on observe que, en changeant $k\theta$ en t, on a

$$(10) \quad \begin{cases} \displaystyle\int_{-h}^{+h} e^{g\theta\sqrt{-1} - k^2\theta^2}\,d\theta = \frac{2}{k}\int_0^{kh} e^{-t^2}\cos\frac{gt}{k}\,dt \\ \displaystyle\qquad\qquad = \frac{2}{k}\int_0^{\infty} e^{-t^2}\cos\frac{gt}{k}\,dt + 2\,\frac{\omega'}{k}, \end{cases}$$

et ω' désigne une quantité évidemment inférieure à $\mathrm{T}(hk)$, et, *a fortiori*, à $\mathrm{T}\left(\frac{hk}{2}\right)$; on pourra donc écrire la formule (8), en ayant égard aux formules (9) et (10), comme il suit :

$$\mathrm{P} = \frac{1}{k\pi}\int_0^{\infty} e^{-t^2}\cos\frac{gt}{k}\,dt + \frac{4\lambda}{\pi k}\,\mathrm{T}\left(\frac{hk}{2}\right) + \frac{\lambda'}{3k^4\pi},$$

en confondant le terme ω' avec celui qui est de même forme. Cette formule peut s'écrire (p. 34)

$$\mathrm{P} = \frac{1}{2k\sqrt{\pi}}\,e^{-\frac{g^2}{4k^2}} + \frac{4\lambda}{\pi k}\,\mathrm{T}\left(\frac{hk}{2}\right) + \frac{\lambda'}{3k^4\pi},$$

ou bien encore, en prenant $h = 1$,

$$P = \frac{1}{\sqrt{2\pi\Sigma pq}} e^{-\frac{g^2}{2\Sigma pq}} + \frac{8\lambda}{\pi\sqrt{2\Sigma pq}} T\left(\frac{\sqrt{\Sigma pq}}{4}\right) + \frac{4\lambda'}{3(\Sigma pq)^2\pi}.$$

Le dernier terme de cette formule est tout à fait négligeable; quant au second, il est beaucoup plus petit que le premier, lorsque g n'est pas un grand nombre.

Si l'on ajoute maintenant les valeurs de P correspondant aux valeurs $i - l$, $i - l + 1$, $i - l + 2, \ldots$, $i + l$ de α, i désignant la différence entre Σp et le plus grand entier N contenu dans Σp, on obtiendra la probabilité Q que E aura lieu un nombre de fois compris entre $N - l$ et $N + l$; on aura donc

$$(11)\quad Q = \sum P = \sum_{i-l}^{i+l} \frac{1}{2k\sqrt{\pi}} e^{-\frac{\alpha^2}{4k^2}} + \frac{8l\lambda}{\pi k} T\left(\frac{k}{2}\right) + \frac{2\lambda' l}{3k^4\pi}.$$

On peut du reste calculer la somme Σ qui entre dans cette équation par la formule d'Euler, et l'on a

$$\sum_{i-l}^{i+l} \frac{1}{2k\sqrt{\pi}} e^{-\frac{g^2}{4k^2}} = \frac{1}{2k\sqrt{\pi}} \int_{-l}^{+l} e^{-\frac{g^2}{4k^2}} dg + \frac{1}{24k^3\sqrt{\pi}} e^{-\frac{l^2}{4k^2}} + \ldots,$$

ce que l'on peut encore écrire

$$\sum \frac{1}{2k\sqrt{\pi}} e^{-\frac{g^2}{4k^2}} = \frac{2}{\sqrt{\pi}} \int_0^{\frac{l}{2k}} e^{-\gamma^2} d\gamma + \frac{l}{24k^3\sqrt{\pi}} e^{-\frac{l^2}{4k^2}} + \ldots$$

En désignant alors par ν une quantité comprise entre 0 et 1, en remplaçant k par sa valeur $\sqrt{\frac{1}{2}\Sigma pq}$, et en se servant de la notation $\Theta(x)$ pour désigner la fonction

$\frac{2}{\sqrt{\pi}}\int_0^x e^{-x^2}dx$, la formule (11) deviendra

$$Q = \Theta\left(\frac{l}{\sqrt{2\Sigma pq}}\right) + l\left[\frac{1}{24k^3\sqrt{\pi}}e^{-\frac{k^2\nu}{4k^2}} + \frac{8\lambda}{\pi k}\mathrm{T}\left(\frac{k}{2}\right) + \frac{2\lambda'}{3k^4\pi}\right].$$

La plupart du temps, si l est de l'ordre de k, mais un peu supérieur à k, on aura à très-peu près $Q = 1$, et l'on pourra poser, sans erreur sensible,

$$Q = \Theta\left(\frac{l}{\sqrt{2\Sigma pq}}\right).$$

Telle est la formule qui donne l'expression de la loi des grands nombres; elle contient évidemment le théorème de Bernoulli. Ce résultat était connu de Poisson, mais personne jusqu'ici n'avait donné l'expression rigoureuse de l'approximation que comportait cette formule.

Théorème inverse de celui de Bernoulli.

Je suppose que l'on ait observé l'arrivée d'un événement E, et que cet événement ait eu lieu α fois sur s épreuves, sa probabilité ne différera pas beaucoup de $\frac{\alpha}{s}$, si le nombre s est très-grand; et, en effet, un événement, dont la probabilité est $\frac{\alpha}{s}$, se produit environ α fois dans s épreuves. Si l'on ne se contente pas de cette approximation un peu vague, on est naturellement conduit à se poser cette question :

Soit p la probabilité constante, mais inconnue, d'un événement E, cet événement ayant été observé α fois dans s épreuves. Quelle est la probabilité que p diffère de $\frac{\alpha}{s}$ d'une quantité inférieure à l.

Soit, comme plus haut, $q = 1 - p$ et $\beta = s - \alpha$. On peut faire sur la valeur de la probabilité p une infinité d'hypothèses qui consistent à lui attribuer successivement toutes les valeurs comprises entre -1 et $+1$. Supposons p compris entre x et $x + dx$. Cette hypothèse particulière donne, *a priori*, à l'événement composé que l'on a observé une probabilité

$$\frac{1.2.3\ldots s}{1.2\ldots\alpha.1.2.3\ldots\beta}x^{\alpha}(1-x)^{\beta},$$

ou infiniment peu différente de cette quantité; et, en vertu du théorème de Bayes (*voir* p. 56), la probabilité que l'événement composé observé est dû à l'hypothèse que nous avons faite est

$$\frac{\frac{1.2.3\ldots s}{\alpha!\,\beta!}x^{\alpha}(1-x)^{\beta}}{\frac{1.2.3\ldots s}{\alpha!\,\beta!}\sum_{0}^{1}x^{\alpha}(1-x)^{\beta}} = \frac{x^{\alpha}(1-x)^{\beta}dx}{\int_{0}^{1}x^{\alpha}(1-x)^{\beta}dx}.$$

Enfin la probabilité P que l'événement observé est dû à l'une des valeurs de p comprises entre $\frac{\alpha}{s} - l$ et $\frac{\alpha}{s} + l$ s'obtiendra en vertu du principe de la probabilité totale, en ajoutant chacune des probabilités telles que la précédente, obtenues en faisant varier x de $+\frac{\alpha}{s} - l$ à $\frac{\alpha}{s} + l$; on a donc enfin

$$P = \frac{\int_{\frac{\alpha}{s}-l}^{\frac{\alpha}{s}+l}x^{\alpha}(1-x)^{\beta}dx}{\int_{0}^{1}x^{\alpha}(1-x)^{\beta}dx},$$

ou bien, en observant que le dénominateur est égal à $\frac{\alpha!\,\beta!}{(s+1)!}$,

$$P = \frac{(s+1)!}{\alpha!\,\beta!}\int_{\frac{\alpha}{s}-l}^{\frac{\alpha}{s}+l} x^{\alpha}(1-x)^{\beta}dx;$$

si l'on pose alors

$$x = \frac{\alpha}{s} + t, \quad (1-x) = \frac{\beta}{s} - t, \quad dx = dt,$$

on trouve

$$P = \frac{\alpha^{\alpha}\beta^{\beta}}{s^{s}}\,\frac{(s+1)!}{\alpha!\,\beta!}\int_{-l}^{+l}\left(1+\frac{st}{\alpha}\right)^{\alpha}\left(1-\frac{st}{\beta}\right)^{\beta}dt.$$

Nous poserons, pour abréger,

(1) $$c = \frac{\alpha^{\alpha}\beta^{\beta}}{s^{s}}\,\frac{(s+1)!}{\alpha!\,\beta!}$$

et nous aurons

(2) $$P = c\int_{-l}^{+l}\left(1+\frac{st}{\alpha}\right)^{\alpha}\left(1-\frac{st}{\beta}\right)^{\beta}dt.$$

Nous supposerons l assez petit pour que $\frac{sl}{\alpha}$ et $\frac{sl}{\beta}$ restent non-seulement inférieurs à l'unité, mais même très-petits et de l'ordre de $\frac{1}{\sqrt{s}}$; on aura alors à peu près

$$\begin{aligned}\log\left(1+\frac{st}{\alpha}\right)^{\alpha}\left(1-\frac{st}{\beta}\right)^{\beta} &= \alpha\left(\frac{st}{\alpha}-\frac{s^2t^2}{2\alpha^2}+\ldots\right)-\beta\left(\frac{st}{\beta}+\frac{s^2t^2}{2\beta^2}+\ldots\right)\\ &= -\frac{s^2t^2}{2}\left(\frac{1}{\alpha}+\frac{1}{\beta}\right) = -\frac{s^3t^2}{2\alpha\beta},\end{aligned}$$

et, par suite, la formule (2) s'écrira

$$P = c\int_{-l}^{+l} e^{-\frac{s^3 t^2}{2\alpha\beta}}\,dt,$$

ou, en changeant de variable et posant $\gamma = l\sqrt{\frac{s^3}{2\alpha\beta}}$,

$$P = c\sqrt{\frac{2\alpha\beta}{s^3}}\int_{-\gamma}^{+\gamma} e^{-\tau^2}\,d\tau.$$

Si l'on évalue c au moyen de la formule approchée de Stirling

$$\alpha! = \alpha^\alpha\sqrt{2\pi\alpha}\,e^{-\alpha},$$

on trouve

$$P = \frac{2}{\sqrt{\pi}}\int_0^\gamma e^{-\tau^2}\,d\tau.$$

Cette formule est rigoureuse quand on suppose $s = \infty$, et il serait facile de calculer l'erreur que nous avons commise, par suite des termes que nous avons négligés; nous nous occuperons tout à l'heure de cette question. Pour le moment, observons que, si l'on suppose $\gamma = \infty$, P devient égal à l'unité; or, pour que γ soit infini, il suffit que $l\sqrt{\frac{s^3}{2\alpha\beta}}$ le soit; mais, si $\frac{\alpha}{s}$ est fini ainsi que $\frac{\beta}{s}$, la quantité $l\sqrt{\frac{s^3}{2\alpha\beta}}$ pourra être infinie, même pour des valeurs infiniment petites de l; on en conclut que la probabilité P, que $\frac{\alpha}{s}$ diffère iufiniment peu de p, est l'unité; donc enfin, si l'on fait un très-grand nombre d'épreuves, $\frac{\alpha}{s}$ tend vers p. C'est le théorème de Bernoulli, démontré en quelque sorte par une méthode inverse de celle que nous avions d'abord suivie pour l'établir.

Essayons maintenant de calculer P avec une approximation plus grande. En nous laissant guider par la méthode précédente, développons $\left(1+\frac{st}{\alpha}\right)^\alpha\left(1-\frac{ts}{\beta}\right)^\beta$ suivant les puissances de t, multiplions ce développement par celui de $e^{\frac{s^2t^2}{2\alpha\beta}}$, nous pourrons écrire

$$(3)\quad \left(1+\frac{st}{\alpha}\right)^\alpha\left(1-\frac{st}{\beta}\right)^\beta e^{\frac{s^2t^2}{2\alpha\beta}} = 1 + k_1 t + k_2 t^2 + k_3 t^3 + \ldots,$$

$k_1, k_2, \ldots$, désignant des coefficients indépendants de t, et, par suite, nous aurons, au lieu de la formule (2),

$$P = c\int_{-l}^{+l}\left(1+\frac{st}{\alpha}\right)^\alpha\left(1-\frac{st}{\beta}\right)^\beta dt$$
$$= c\int_{-l}^{+l} e^{-\frac{s^2t^2}{2\alpha\beta}}(1 + k_1 t + k_2 t^2 + \ldots)dt;$$

les termes en $t, t^3, t^5, \ldots$ disparaissent d'eux-mêmes dans cette formule, et l'on peut se contenter d'écrire

$$P = 2c\int_0^l e^{-\frac{s^2t^2}{2\alpha\beta}}(1 + k_2 t^2 + k_4 t^4 + \ldots)dt,$$

et, en posant toujours

$$\gamma = l\sqrt{\frac{s^3}{2\alpha\beta}},$$

on aura

$$P = 2c\sqrt{\frac{2\alpha\beta}{s^3}}\int_0^\gamma e^{-\gamma^2}\left[1 + k_2\gamma^2\frac{2\alpha\beta}{s^3} + k_4\gamma^4\left(\frac{2\alpha\beta}{s^3}\right)^2 + \ldots\right]d\gamma.$$

Le facteur en dehors du signe $\int$ est à peu près égal à $\frac{2}{\sqrt{\pi}}$;

on peut du reste calculer sa valeur exacte par la formule de Stirling; nous le représenterons par $\frac{2\omega}{\sqrt{\pi}}$. Si nous posons ensuite

$$\Theta_i(\gamma) = \int_0^\gamma e^{-\gamma^2}\gamma^i d\gamma,$$

nous aurons

$$(4)\quad P = \frac{2\omega}{\sqrt{\pi}}\left[\Theta_0(\gamma) + k_2\frac{2\alpha\beta}{s^3}\Theta_2(\gamma) + k_4\left(\frac{2\alpha\beta}{s^3}\right)^2\Theta_4(\gamma) + \ldots\right].$$

Chacune des intégrales Θ_i se calcule aisément en fonction de Θ_0 au moyen d'intégrations par parties; il ne reste donc plus, pour résoudre la question, qu'à calculer les coefficients k_2, k_4, ... Revenons pour cela à la formule (3). En prenant les différentielles logarithmiques des deux membres par rapport à t, on aura

$$(5)\quad \frac{s}{1+\frac{st}{\alpha}} - \frac{s}{1-\frac{st}{\beta}} + \frac{s^3 t}{\alpha\beta} = \frac{k_1 + 2k_2 t + k_3 t^2 + \ldots}{1 + k_1 t + k_2 t^2 + \ldots};$$

en développant le premier membre suivant les puissances de t, on trouve

$$s\left(1 - \frac{st}{\alpha} + \frac{s^2t^2}{\alpha^2} - \ldots\right) - s\left(1 + \frac{st}{\beta} + \frac{s^2t^2}{\beta^2} + \ldots\right) + \frac{s^3 t}{\alpha\beta},$$

ou, en réduisant,

$$-s^3t^2\left(\frac{1}{\beta^2} - \frac{1}{\alpha^2}\right) - s^4t^3\left(\frac{1}{\beta^3} + \frac{1}{\alpha^3}\right) - \ldots;$$

la formule (5) pourra donc s'écrire

$$-\left[s^3 t^2\left(\frac{1}{\beta^2}-\frac{1}{\alpha^2}\right)+s^4 t^3\left(\frac{1}{\beta^3}+\frac{1}{\alpha^3}\right)\right.$$
$$\left.+s^5 t^4\left(\frac{1}{\beta^4}-\frac{1}{\alpha^4}\right)+\ldots\right](1+k_1 t+k_2 t^2+\ldots)$$
$$=k_1+2k_2 t+3k_3 t^2+\ldots;$$

en égalant alors les coefficients des mêmes puissances de t, on a

$$k_1=0,\quad k_2=0,$$
$$3k_3=-s^3\left(\frac{1}{\beta^2}-\frac{1}{\alpha^2}\right),$$
$$4k_4=-s^3\left(\frac{1}{\beta^2}-\frac{1}{\alpha^2}\right)k_1-s^4\left(\frac{1}{\beta^3}+\frac{1}{\alpha^3}\right),$$
$$5k_5=-s^3\left(\frac{1}{\beta^2}-\frac{1}{\alpha^2}\right)k_2-s^4\left(\frac{1}{\beta^3}+\frac{1}{\alpha^3}\right)k_1-s^5\left(\frac{1}{\beta^4}-\frac{1}{\alpha^4}\right),$$
$$\ldots\ldots\ldots,$$

c'est-à-dire

$$k_2=0,\quad k_4=-\frac{s^4}{4}\left(\frac{1}{\beta^3}+\frac{1}{\alpha^3}\right)\cdots;$$

on aura ainsi, au lieu de la formule (4),

$$(6)\qquad P=\frac{2\omega}{\sqrt{\pi}}\left[\Theta_0(\gamma)-\frac{\alpha^2\beta^2}{s^2}\left(\frac{1}{\beta^3}+\frac{1}{\alpha^3}\right)\Theta_4(\gamma)-\ldots\right].$$

Évaluation du reste dans la série donnée au paragraphe précédent.

Nous avons trouvé les coefficients k_1, k_2, ... par la voie récurrente; on aurait pu en obtenir l'expression générale, mais sous une forme compliquée et qui n'aurait rien appris sur leur valeur numérique, en effectuant le produit des

trois fonctions $\left(1+\frac{st}{\alpha}\right)^\alpha$, $\left(1-\frac{st}{\beta}\right)^\beta$ et $e^{\frac{s^2t^2}{2\alpha\beta}}$ directement. La voie récurrente et la méthode directe ne nous donnent ni l'une ni l'autre des formules assez simples pour estimer une limite de l'erreur commise en s'arrêtant à un terme déterminé; il était cependant important d'évaluer cette erreur, afin de fixer rigoureusement le degré d'exactitude des formules du paragraphe précédent : c'est ce qui n'avait pas encore été fait jusqu'ici.

Je vais commencer par évaluer une limite supérieure de la valeur du coefficient k_n. En posant

$$f(t)=\left(1+\frac{st}{\alpha}\right)^\alpha\left(1-\frac{st}{\beta}\right)^\beta e^{\frac{s^2t^2}{2\alpha\beta}},$$

on a

$$\log f(t)=\alpha\log\left(1+\frac{st}{\alpha}\right)+\beta\log\left(1-\frac{st}{\beta}\right)+\frac{s^2t^2}{2\alpha\beta}$$

ou

$$(1)\quad \log f(t)=-\frac{s^3t^3}{3}\left(\frac{1}{\beta^2}-\frac{1}{\alpha^2}\right)-\frac{s^4t^4}{4}\left(\frac{1}{\beta^3}+\frac{1}{\alpha^3}\right)-\ldots,$$

d'où l'on tire, en repassant des logarithmes aux nombres,

$$f(t)=1-\left[\frac{s^3t^3}{3}\left(\frac{1}{\beta^2}-\frac{1}{\alpha^2}\right)+\frac{s^4t^4}{4}\left(\frac{1}{\beta^3}+\frac{1}{\alpha^3}\right)\cdots\right]$$
$$+\frac{1}{1.2}\left[\frac{s^3t^3}{3}\left(\frac{1}{\beta^2}-\frac{1}{\alpha^2}\right)+\ldots\right]^2$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

On ne fera qu'augmenter le coefficient k_n de t^n si, au lieu de $f(t)$, on considère la fonction suivante, où δ est le plus

petit des deux nombres α, β,

$$1 + \frac{s^3 t^3}{\delta^2} + \frac{s^4 t^4}{\delta^3} + \ldots + \frac{1}{1.2}\left(\frac{s^3 t^3}{\delta^2} + \ldots\right)^2 + \frac{1}{1.2.3}\left(\frac{s^3 t^3}{\delta^2} + \ldots\right)^3 + \ldots,$$

ou bien

$$1 + \frac{s^3 t^3}{\delta^2 - \delta st} + \frac{1}{1.2}\left(\frac{s^3 t^3}{\delta^2 - \delta st}\right)^2 + \frac{1}{1.2.3}\left(\frac{s^3 t^3}{\delta^2 - \delta st}\right)^3 + \ldots.$$

En développant par la formule du binôme chacun des termes de cette expression et en écrivant que le coefficient de t^n est plus grand que k_n, on aura

$$(2)\quad \left\{ \begin{aligned} k_n < s^n \Bigg[& \frac{1}{\delta^{n-1}} + \frac{n-5}{1.2} \frac{1}{\delta^{n-2}} + \frac{(n-6)(n-7)}{1.2.2.3} \frac{1}{\delta^{n-3}} \\ & + \frac{(n-7)(n-8)(n-9)}{1.2.2.3.3.4} \frac{1}{\delta^{n-4}} + \ldots \Bigg]. \end{aligned} \right.$$

Le dernier terme du second membre a pour dénominateur $\delta^{\mu(n)}$, le nombre $\mu(n)$ représentant deux fois le plus grand entier contenu dans $\frac{n}{3}$, augmenté du reste de la division de n par 3; or, en écrivant comme il suit la formule (2)

$$\begin{aligned} k_n < \frac{2s^n}{\delta^{n-1}} \Bigg(& \frac{1}{2} + \frac{\frac{n}{4} - \frac{5}{4}}{1}\delta + \frac{\frac{n}{4} - \frac{6}{4}}{1} \frac{\frac{n}{3} - \frac{7}{3}}{2}\delta^2 \\ & + \frac{\frac{n}{4} - \frac{7}{4}}{1} \frac{\frac{n}{3} - \frac{8}{3}}{2} \frac{\frac{n}{4} - \frac{9}{4}}{3}\delta^3 + \ldots \Bigg), \end{aligned}$$

on reconnaît immédiatement que si nous désignons par ν la quantité $n - 1 - \mu(n)$, $\mu(n)$ étant toujours égal ou plus

petit que $\frac{2n}{3}+1$, nous aurons

$$\nu > n - \frac{2n}{3} - 2 \quad \text{ou} \quad \nu > \frac{n}{3} - 2,$$

et l'on pourra écrire

$$k_n < \frac{2s^n}{\delta^{n-1}}\left[1 + \frac{\nu}{1}\delta + \frac{\nu(\nu-1)}{1.2}\delta^2 + \ldots + \delta^\nu\right],$$

ou bien

$$k_n < \frac{2s^n}{\delta^{n-1}}(1+\delta)^\nu \quad \text{ou} \quad < \frac{2s^n}{\delta^{\mu(n)}}\frac{(1+\delta)^\nu}{\delta^\nu} \quad \text{ou} \quad < \frac{2s^n}{\delta^{\mu(n)}}\left(1+\frac{1}{\delta}\right)^\nu.$$

Reprenant alors la formule (4) du paragraphe précédent,

$$P = \frac{2\omega}{\sqrt{\pi}}\left[\Theta_0(\gamma) + k_4\Theta_4(\gamma)\left(\frac{2\alpha\beta}{s^3}\right)^2 + \ldots\right],$$

la connaissance de la limite supérieure de k_n permettra de calculer l'erreur commise en négligeant le reste de la série à partir d'un terme quelconque; nous examinerons particulièrement le cas où l'on calcule P, en se bornant au premier terme de la série. En appelant ε l'erreur, on a

$$\varepsilon < \frac{4\omega}{\sqrt{\pi}}\left[\left(\frac{2\alpha\beta}{s}\right)^2\frac{1}{\delta^3}\Theta_4(\gamma) + \left(\frac{2\alpha\beta}{s}\right)^3\frac{1}{\delta^4}\left(1+\frac{1}{\delta}\right)\Theta_6(\gamma) + \ldots\right].$$

Il est à remarquer que $\mu(n)$ croît de quatre unités quand n croît de trois unités, et, par suite, le second membre de la formule précédente décroît plus rapidement qu'une progression géométrique; enfin on peut remplacer $\Theta_4(\gamma)$, $\Theta_6(\gamma), \ldots$ par $\frac{\gamma^5}{5}, \frac{\gamma^7}{7}, \ldots$, et l'on a

$$\varepsilon < \frac{4\omega}{\sqrt{\pi}}\left[\left(\frac{2\alpha\beta}{s}\right)^2\frac{1}{\delta^3}\frac{\gamma^5}{5} + \left(\frac{2\alpha\beta}{s}\right)^3\frac{1}{\delta^4}\left(1+\frac{1}{\delta}\right)\frac{\gamma^7}{7} + \ldots\right.$$

Ces deux termes suffiront pour calculer ε; le suivant contient non pas $\frac{1}{\delta^5}$, mais $\frac{1}{\delta^6}$ en facteur, ainsi qu'on peut s'en assurer. On sommera ce reste comme une progression géométrique ayant pour raison $\frac{2\alpha\beta}{s\delta}\gamma^2\frac{1}{\delta}$, ce qui est permis en groupant les termes deux à deux quand les exposants de δ vont en croissant d'une unité seulement, et en doublant les termes pour lesquels les exposants de δ décroissent de deux unités. Il y aura aussi parfois avantage à remplacer $\Theta_4(\gamma)$, $\Theta_6(\gamma)$,... par les valeurs $\Theta_4(\infty)$, $\Theta_6(\infty)$,..., qui sont égales à $\frac{\sqrt{\pi}}{2}\frac{3}{4}$, $\frac{\sqrt{\pi}}{2}\frac{15}{8}$, ...

Nous ne simplifierons pas la valeur de ε et nous préférons laisser le reste sous la forme d'une série, parce que, selon les valeurs de α, β, s, il conviendra de modifier l'usage de cette série.

Il est presque inutile de faire observer que la série en question peut servir à calculer une limite de l'erreur commise dans l'évaluation de P, quand au lieu de négliger le terme en Θ_4 on néglige un terme d'ordre quelconque, et le reste de la série qui donne P est toujours inférieur au reste de la série qui fait connaître la limite de l'erreur ε.

Le reste est généralement assez considérable pour qu'il y ait lieu de calculer les termes en Θ_4 et en Θ_6, au moyen des formules

$$\Theta_2 = \int_0^\gamma e^{-\gamma^2}\gamma^2 d\gamma = -\frac{\gamma}{2}e^{-\gamma^2} + \frac{1}{2}\Theta_0(\gamma),$$

$$\Theta_4 = \int_0^\gamma e^{-\gamma^2}\gamma^4 d\gamma = -\frac{\gamma^3}{2}e^{-\gamma^2} + \frac{3}{2}\Theta_2(\gamma),$$

$$\Theta_6 = \int_0^\gamma e^{-\gamma^2}\gamma^6 d\gamma = -\frac{\gamma^5}{2}e^{-\gamma^2} + \frac{5}{2}\Theta_4(\gamma),$$

....

Pour faire une application, nous supposerons que l'événement E ait été observé 8000 fois sur 10000 épreuves. La probabilité de E sera $\frac{8000}{10000}$ ou $\frac{4}{5}$, et il y aura une probabilité P que l'erreur est inférieure à l, P étant donné par la formule

$$P = \frac{2\omega}{\sqrt{\pi}} \int_0^{2,4,\dots} e^{-\gamma^2} d\gamma = \omega . 0,999;$$

mais alors on aura

$$2,4 = l \sqrt{\frac{(10\,000)^3}{2.8000.2000}},$$

d'où l'on conclura

$$l = 0,013\dots.$$

Ainsi il y aura environ mille à parier contre un que la probabilité de E est $\frac{4}{5}$ plus ou moins un centième, et l'erreur commise dans l'évaluation de P sera inférieure à 0,06. Ce chiffre est trop fort; il conviendra alors de prendre γ plus petit; si l'on prend, par exemple, $\gamma = 1,8$, on aura seulement $P = 0,99$, et l'erreur commise sur P ne sera plus guère que de 0,01.

Solutions de quelques questions relatives aux épreuves répétées.

PROBLÈME I. — *Une première urne contient 1 boule noire et a boules blanches, une deuxième urne contient 1 boule noire et b boules blanches, une troisième urne contient 1 boule noire et c boules blanches, etc. On tire une boule au hasard dans la première urne; si elle n'est pas blanche, on en tire une dans la deuxième; si elle n'est*

pas blanche, on en tire une dans la troisième, et ainsi de suite ; on demande la probabilité que l'on a de tirer une boule blanche, en accordant : 1° une seule épreuve, 2° deux épreuves, 3° trois épreuves, etc.

1° La probabilité que l'on a de tirer une boule blanche du premier coup est $\frac{a}{a+1}$. 2° La probabilité que l'on a de tirer une boule blanche en accordant deux coups est une probabilité totale égale à la probabilité de tirer une boule blanche du premier coup $\frac{a}{a+1}$, plus la probabilité $\frac{1}{a+1}$ de ne pas tirer une boule blanche du premier coup et d'en tirer une au second; cette probabilité est $\frac{1}{a+1}\,\frac{b}{b+1}$; ainsi la probabilité de tirer une boule blanche, du premier coup ou au second coup seulement, est

$$\frac{a}{a+1}+\frac{b}{(a+1)(b+1)}.$$

3° La probabilité de tirer une boule blanche, quand on accorde la permission d'effectuer trois épreuves, est égale à la probabilité $\frac{a}{a+1}$ de tirer une boule blanche du premier coup, plus la probabilité $\frac{b}{(a+1)(b+1)}$ de tirer une boule noire au premier coup et une boule blanche au second, plus la probabilité $\frac{c}{(a+1)(b+1)(c+1)}$ de tirer une boule noire au premier et au second coup, mais une blanche au troisième ; cette probabilité est donc

$$\frac{a}{a+1}+\frac{b}{(a+1)(b+1)}+\frac{c}{(a+1)(b+1)(c+1)},$$

et ainsi de suite. Si l'on permet d'effectuer un nombre illimité d'épreuves, il est clair que l'on finira par tirer une boule blanche, et, par suite, on a identiquement

$$1 = \frac{a}{a+1} + \frac{b}{(a+1)(b+1)} + \frac{c}{(a+1)(b+1)(c+1)} + \ldots.$$

Cette identité, qui a cela de remarquable, qu'elle contient une infinité d'arbitraires, peut se démontrer directement, avec la plus grande facilité, en prenant, pour a, b, $c, \ldots$, des nombres positifs, tels que le produit

$$(a+1)(b+1)(c+1)\ldots$$

croisse au delà de toute limite. On démontrerait avec une égale facilité la formule

$$1 = \frac{a}{a+a'} + \frac{a'b}{(a+a')(b+b')} + \frac{a'b'c}{(a+a')(b+b')(c+c')} + \ldots,$$

soit directement, soit à l'aide du calcul des probabilités.

Si l'on prend $a = a' = b = b' = \ldots = 1$, on a la formule connue

$$1 = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots;$$

pour $a' = b' = c' = \ldots = 1$ et $a = 1$, $b = 2$, $c = 3, \ldots$, on trouve

$$1 = \frac{1}{2} + \frac{1}{3} + \frac{1}{2.4} + \frac{1}{2.3.5} + \frac{1}{2.3.4.6} + \ldots,$$

ce que l'on peut vérifier, en observant que le second membre est égal à

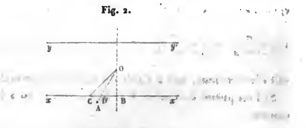
$$\int_0^1 \left(x + \frac{x^2}{1} + \frac{x^3}{1.2} + \frac{x^4}{1.2.3} + \ldots\right) dx = \int_0^1 x e^x dx = 1.$$

Dans le courant de cet Ouvrage, nous aurons souvent l'occasion de nous appuyer sur le théorème de Bernoulli; nous croyons cependant devoir en présenter, dès à présent, une application qui avait déjà attiré l'attention de Laplace et qui a donné lieu depuis à divers travaux intéressants.

PROBLÈME II. — *On donne une série de droites parallèles, équidistantes, tracées sur un plancher très-uni; sur ce plancher on projette une aiguille très-fine; on demande la probabilité que cette aiguille rencontre l'une des parallèles, lorsqu'elle sera parvenue au repos.*

Soient $2a$ la distance de deux des parallèles xx' et yy' quelconques, $2l$ la longueur de l'aiguille. Il est clair que

Fig. 2.



l'on ne change pas le rapport des cas favorables et des cas possibles : 1° en ne considérant qu'une moitié de l'aiguille; 2° en supposant que le centre O tombe sur une perpendiculaire donnée à xx' et à une distance de xx' au plus égale à a; 3° en supposant que la moitié considérée tombe à gauche de OB, en faisant avec OB un angle au plus égal à $\frac{\pi}{2}$. Nous ferons alors ces hypothèses, ce qui ne change en rien la probabilité de l'événement attendu.

Appliquons le principe de la probabilité totale, en examinant ce qui se passe dans chacune des hypothèses où le centre de l'aiguille tombe sur chacun des éléments dx de la

perpendiculaire OB, comptée depuis son pied B jusqu'à la moitié de l'intervalle $2a$ qui sépare les parallèles. La probabilité de chacune de ces hypothèses est $\frac{dx}{a}$; l'une d'elles se trouvant réalisée et l'aiguille tombant dans une position telle que OA, la probabilité qu'il y ait rencontre est le rapport de l'angle COB à $\frac{\pi}{2}$, OC désignant la position que prend l'aiguille quand sa pointe vient raser la parallèle xx'; mais, OB étant désigné par x, l'angle COB a pour cosinus $\frac{x}{l}$; la probabilité de la rencontre, dans l'hypothèse particulière où nous nous sommes placés, est donc

$$\frac{2}{\pi} \operatorname{arc} \cos \frac{x}{l},$$

et la probabilité totale est

$$P = \int_0^l \frac{2}{\pi} \operatorname{arc} \cos \frac{x}{l} \frac{dx}{a},$$

ou bien

$$P = \frac{2}{a\pi} \int_0^l \operatorname{arc} \cos \frac{x}{l}\, dx;$$

en effectuant l'intégration par parties, on trouve

$$P = \frac{2}{a\pi} \int_0^l \frac{x\,dx}{\sqrt{l^2 - x^2}},$$

ou bien

$$P = \frac{2l}{a\pi}.$$

On arrive encore à ce résultat d'une autre manière; en effet, supposons qu'au lieu d'une aiguille on projette un

disque convexe, assez petit pour ne pas rencontrer deux parallèles à la fois, et de périmètre S; chacun de ses éléments ds a la même probabilité de rencontrer les parallèles; en désignant par pds cette probabilité, la probabilité que le disque rencontrera les parallèles sera $\Sigma p ds$ ou $p\frac{S}{2}$: j'écris $\frac{pS}{2}$ et non pS, parce que la rencontre du disque a toujours lieu par deux éléments, ce qui diminue de moitié le nombre des cas favorables à considérer; pour calculer p, on peut faire une application au cercle de rayon R; la probabilité que ce cercle rencontrera est évidemment égale au rapport du nombre des cas où son centre sera à une distance inférieure à R de la parallèle la plus rapprochée à celui des cas où ce centre est à une distance inférieure à a de cette même parallèle; en d'autres termes, pour le disque circulaire, la probabilité cherchée est $\frac{R}{a}$. On a, du reste,

$$S = 2\pi R;$$

on a donc, pour déterminer p, la formule

$$\frac{1}{2} p\, 2\pi R = \frac{R}{a}, \quad \text{d'où} \quad p = \frac{1}{\pi a};$$

il en résulte que $\frac{S}{2a\pi}$ est la probabilité relative à un disque convexe quelconque. (Si le disque n'était pas convexe, on pourrait évidemment le remplacer par un autre convexe, obtenu en menant toutes les tangentes, communes et extérieures, à deux portions du périmètre.)

Une aiguille de longueur $2l$ peut être considérée comme un disque convexe de longueur $4l$ et, par suite, la probabilité relative à l'aiguille sera $\frac{2l}{a\pi}$.

L'application du théorème de Bernoulli à la question que nous venons de résoudre conduit à cette conséquence curieuse que, si on laisse tomber une aiguille, de longueur l, sur une série de lignes parallèles, tracées sur un plancher, distantes entre elles d'une quantité a, le rapport du nombre total des épreuves au nombre des cas où la rencontre de l'aiguille avec les lignes aura lieu tendra sans cesse vers $\frac{a\pi}{2l}$; connaissant a et l, on en déduit π. Si $a = 2l$, on trouve π lui-même en effectuant la division.

Voyons maintenant sur quelle approximation on peut compter en appliquant ce procédé; nous avons vu que la probabilité que le rapport du nombre des cas où l'événement attendu a lieu au nombre total s des épreuves diffère de la probabilité de cet événement d'une quantité inférieure à $\frac{\varepsilon}{s}$, avec une probabilité

$$\frac{2}{\sqrt{\pi}}\int_0^{\frac{\varepsilon}{\sqrt{2pqs}}} e^{-\gamma^2}d\gamma.$$

Si l'on prend $\frac{\varepsilon}{\sqrt{2pqs}}$ égal à $2,4$, l'intégrale précédente devient égale à $0,9999$; posons donc

$$\frac{\varepsilon}{\sqrt{2pqs}} = 2,4.$$

Remplaçons p par $\frac{2l}{a\pi}$, q par $1-\frac{2l}{a\pi}$, nous aurons

$$\frac{\varepsilon^2}{s^2} = (2,4)^2 2\left(1-\frac{2l}{a\pi}\right)\frac{2l}{a\pi}\frac{1}{s};$$

$\frac{\varepsilon^2}{s^2}$ est le carré de l'erreur commise en appliquant notre pro-

cédé au calcul de π; donc l'erreur variecomme l'inverse de la racine carrée du nombre des épreuves et, toutes choses égales d'ailleurs, atteindra son maximum pour $\frac{2l}{a\pi} = \frac{1}{2}$, où $l = \frac{a\pi}{4}$. Prenons, par exemple, $l = a$, on aura

$$\frac{\varepsilon}{s} < \sqrt{\frac{1}{2s}};$$

si l'on prend $s = 72$, on aura

$$\frac{\varepsilon}{s} < \frac{1}{12};$$

pour $s = 5000$, on a

$$\frac{\varepsilon}{s} < \frac{1}{100}.$$

On voit donc que le procédé que nous proposons est, en général, peu sensible et plus curieux qu'utile. Il est vrai que, si $\frac{2l}{a\pi}$ était très-petit, on pourrait espérer une approximation plus grande, mais la formule qui donne l'approximation cesserait d'être exacte.

Espérance. — Importance d'une somme éventuelle.

Lorsqu'une personne attend une somme d'argent qui dépend de l'arrivée d'un événement incertain, on dit que cette somme est *éventuelle.*

L'importance d'une somme éventuelle peut être mesurée avec précision, et voici comment. Supposons d'abord que cette somme soit a et qu'elle dépende de l'arrivée de l'événement E (on peut supposer, par exemple, que l'événe-

ment E soit le tirage d'un numéro déterminé d'une loterie; si l'on tire ce numéro, on gagne la somme a; si l'on en tire un autre, on ne réalise aucun bénéfice); soit p la probabilité de E. Si l'on fait un grand nombre s d'épreuves, en vertu du théorème de Bernoulli, l'événement E se présentera un nombre de fois α, l'événement contraire se présentera β fois et le bénéfice réalisé sera $a\alpha$; mais on a, à fort peu de chose près,

$$\frac{\alpha}{s} = p,$$

et, pour $s = \infty$, cette formule devient rigoureuse; on a donc

$$\frac{a\alpha}{s} = ap;$$

ap est donc le bénéfice moyen et ap diffère infiniment peu du prix qu'il faudrait payer le droit de faire une épreuve, pour ne faire ni perte ni gain dans une longue série d'épreuves.

Il y a plus; si nous posons

$$q = 1 - p \quad \text{et} \quad \Theta(\gamma) = \frac{2}{\sqrt{\pi}} \int_0^{\gamma} e^{-\gamma^2} d\gamma \quad (voir \text{ p. 97 et 106}),$$

il y aura une probabilité

$$P = \Theta\left(\frac{l}{\sqrt{2pqs}}\right)$$

que l'événement E aura lieu un nombre de fois compris entre $\alpha - l$ et $\alpha + l$ sur s épreuves, α étant à peu près ps. Donc il y aura la probabilité P que le gain réalisé sera compris entre

$$(\alpha + l)a \quad \text{et} \quad (\alpha - l)a,$$

et, si l'on paye $ap \pm \varepsilon$ le droit de faire une épreuve, comme on a déboursé en tout $sap \pm s\varepsilon$ ou $a\alpha \pm s\varepsilon$, le bénéfice net aura la probabilité P d'être compris entre $al \pm \varepsilon s$ et $-al \pm \varepsilon s$. Si l'on prend l d'un ordre un peu supérieur à $\sqrt{s}$, al sera négligeable vis-à-vis de εs, et P sera très-voisin de l'unité; il sera donc très-probable que l'on réalisera un bénéfice net $\pm \varepsilon s$; en d'autres termes, ap mesure l'*importance* de la somme éventuelle a, puisque, en payant $ap + \varepsilon$, on finira par être toujours en perte, et en payant $ap - \varepsilon$ on finira par être toujours en gain.

On appelle *espérance mathématique* d'une somme a le produit de cette somme par la probabilité p de l'obtenir dans une épreuve. On peut donc énoncer le théorème suivant :

Théorème I. — *L'importance d'une somme éventuelle est égale à son espérance mathématique.*

Supposons maintenant que la valeur absolue d'une somme éventuelle dépende de la manière dont l'événement E se présente (supposons, pour fixer les idées, qu'un coup de dé procure le bénéfice 1, 2, 3,..., 6 francs, selon que l'on amène le point 1, 2,..., 6); la valeur de cette somme, ou le prix que l'on pourra payer le droit d'effectuer l'épreuve, sera évidemment la somme des prix que l'on devra payer chacune des manières dont l'événement peut se présenter; ainsi, dans l'exemple choisi, il faudra payer $\frac{1^f}{6}$ le droit d'attendre le bénéfice procuré par le point as, $\frac{2^f}{6}$ le droit d'attendre le bénéfice dû au point 2, etc. Nous pouvons donc énoncer le théorème suivant :

Théorème II. — *Lorsque la valeur absolue d'une somme éventuelle varie avec la manière dont l'événe-*

ment E, *dont elle dépend, se présente, l'importance de cette somme est mesurée par la somme des espérances mathématiques, relatives à chacune des façons dont* E *peut se présenter.*

Des jeux de hasard.

Bien que tout le monde sache parfaitement ce que l'on appelle un *jeu de hasard,* je crois qu'il est bon de donner une définition précise de ce que nous devrons entendre par ces mots; pour nous, un jeu sera une opération quelconque dans laquelle un ou plusieurs individus seront assujettis à attendre un bénéfice éventuel, positif ou négatif; ces individus porteront le nom de *joueurs*. Cette définition du jeu nous permettra de traiter de la même façon un grand nombre de questions importantes et, en particulier, la théorie des assurances sur la vie humaine et sur les choses.

Dans la plupart des cas, les joueurs sont tenus de payer l'avantage que leur procure le droit d'attendre les sommes éventuelles qui constituent le gain du jeu, et la première question dont nous allons nous occuper est de régler équitablement la mise des joueurs.

Pour qu'un jeu ne soit ni avantageux ni désavantageux, il faut évidemment que, dans un grand nombre de parties ou d'épreuves, le gain ou la perte moyenne de chaque joueur soit sensiblement zéro, et que, si l'on altère infiniment peu les mises, le gain ou la perte moyenne devienne nécessairement différent de zéro.

Cela revient à dire que *la mise de chaque joueur doit être égale à l'importance de la somme éventuelle qu'il attend.* Supposons, ce qui est le cas le plus ordinaire, que la somme éventuelle attendue par chaque joueur soit la

mise totale des joueurs; soit a_i la mise d'un quelconque des joueurs, p_i la probabilité qu'il a de gagner et $q_i = 1 - p_i$ la probabilité qu'il a de perdre; la valeur de la somme qu'il attend est $p_i \Sigma a$. On doit donc avoir

$$a_i = p_i \Sigma a \quad \text{ou} \quad \frac{a_i}{p_i} = \Sigma a,$$

d'où l'on conclut

$$\frac{a_1}{p_1} = \frac{a_2}{p_2} = \ldots = \frac{a_i}{p_i} = \ldots,$$

et, dans ce cas, la mise de chaque joueur doit être proportionnelle à la chance qu'il a de gagner.

Les conditions du jeu ne sont pas toujours aussi simples, et la théorie dont nous nous occupons dans ce Chapitre aura surtout pour but de régir équitablement les mises des joueurs, mais toujours d'après le même principe, à savoir que *la mise doit être égale à l'importance de la somme attendue*.

Sur l'effet des jeux de hasard.

Le Calcul des probabilités, d'accord avec les préceptes de la morale, qui, considérée indépendamment de toute idée religieuse, n'est que l'application des règles du bon sens à notre conduite, nous engage à modérer en nous la passion du jeu; nous allons prouver, en effet, que les habitudes du jeu conduisent fatalement à la ruine.

Considérons un individu A ayant l'habitude du jeu; son sort ne sera pas modifié si l'on substitue à ses adversaires un joueur unique et idéal capable d'acheter constamment leurs profits et pertes; j'appellerai M ce joueur idéal, dont la fortune peut être censée aussi grande que l'on veut, puis-

qu'il représente le public. Ceci posé, je vais prouver que, si A et M jouent ensemble un jeu équitable, le plus riche des deux joueurs ruinera probablement l'autre.

Soient a la fortune de A et m celle de M, ces fortunes peuvent être considérées comme des mises et, si l'on appelle x et y les probabilités que A et M ont de gagner dans un jeu équitable, il faut que l'importance de la mise m pour A soit la même que l'importance de la mise a pour M : donc mx doit être égal à ay. On a donc

$$mx = ay, \quad x + y = 1,$$

d'où l'on tire

$$x = \frac{a}{a+m}, \quad y = \frac{m}{a+m}.$$

Si M représente le public, on a

$$m = \infty \quad \text{et, par suite,} \quad x = 0, \quad y = 1$$

et la probabilité y que A a de se ruiner devient une certitude (comparer ce résultat avec celui de la p. 79).

Il en est de cette vérité comme du mouvement perpétuel en Mécanique. Il est aussi difficile de faire comprendre à un joueur qu'il court à sa ruine qu'il est difficile de démontrer aux gens dont l'éducation scientifique a été négligée l'inanité de leurs combinaisons, lorsqu'ils veulent faire des machines créant indéfiniment du travail utile. Les joueurs ont leurs superstitions, comme ceux qui croient au surnaturel, et il est remarquable que des personnes, d'ailleurs intelligentes, refusent, dans certains cas, de vouloir faire usage de leur raison pour accepter, sans examen, des doctrines, tout au moins ridicules lorsqu'elles ne sont point funestes. Nous écrivons ici pour des gens raisonnables et nous ne nous appliquerons point à rétorquer tous les arguments, à détruire tous les préjugés des joueurs ; ainsi que nous l'avons

dit, il existe des gens dont un côté de la raison est resté atrophié et pour lesquels la foi domine les arguments les plus sûrs et les plus solides.

Solution de quelques questions pour montrer l'usage que l'on peut faire de l'espérance mathématique.

Problème I. — *Deux joueurs* A *et* B *jouent aux conditions suivantes : celui qui gagne une partie marque un point, dès que* A *a marqué* α *points ou que* B *en a marqué* β, *le jeu cesse, et* A *gagne la mise de* B *dans le premier cas,* B *gagne la mise de* A *dans le second ; quelles doivent être les mises de* A *et* B *pour que le jeu soit équitable ?*

Calculons la probabilité que chaque joueur a de gagner ; pour cela, désignons par $f(u, v)$ la probabilité que, sur $u+v$ parties jouées, A en gagnera u et B en gagnera v. On a, comme l'on sait (p. 88),

$$(1) \qquad f(u, v) = \frac{(u+v)!}{u!\,v!}\left(\frac{1}{2}\right)^{u+v}.$$

A peut gagner la mise de B : 1° en marquant α points de suite dans les α premières parties ou, ce qui revient au même, $\alpha-1$ points dans les $\alpha-1$ premières parties et 1 point dans la $\alpha^{\text{ième}}$; 2° en marquant $\alpha-1$ points dans les α premières parties et 1 dans la $\alpha+1^{\text{ième}}$; 3° en marquant $(\alpha-1)$ points dans les $\alpha+1$ premières parties et 1 dans la $\alpha+2^{\text{ième}}$, etc., jusqu'à la $\alpha+\beta-1^{\text{ième}}$ partie. La probabilité p que A a de gagner la mise de B s'obtiendra donc en appliquant le principe de la probabilité totale, les probabilités des causes successives étant les moitiés de

$$f(\alpha-1, 0), \quad f(\alpha-1, 1), \ldots, \quad f(\alpha-1, \beta-1)$$

et les probabilités qu'elles donnent à l'événement attendu étant 1, on trouvera ainsi

$$p=\frac{1}{2}[f(\alpha-1,\ 0)+f(\alpha-1,\ 1)+\ldots+f(\alpha-1,\ \beta-1)],$$

c'est-à-dire, en vertu de (1),

$$p=\frac{1}{2}\,\frac{1}{(\alpha-1)!}\left[\frac{(\alpha-1)!}{2^{\alpha-1}}+\frac{\alpha!}{1!\,2^{\alpha}}\right.$$
$$\left.+\frac{(\alpha+1)!}{2!\,2^{\alpha+1}}+\ldots+\frac{(\alpha+\beta-2)!}{(\beta-1)!\,2^{\alpha+\beta-2}}\right].$$

On trouverait de même

$$1-p=\frac{1}{2}\,\frac{1}{(\beta-1)!}\left[\frac{(\beta-1)!}{2^{\beta-1}}+\frac{\beta!}{1!\,2^{\beta}}+\ldots+\frac{\alpha+\beta-2}{(\alpha-1)!\,2^{\alpha+\beta-2}}\right];$$

l'élimination de p entre ces deux formules conduit à une identité qu'il serait plus difficile d'établir directement; du reste, p et $1-p$ sont proportionnels aux mises de A et B : la question se trouve donc ainsi résolue.

Si l'on fait $\alpha=\beta$, on trouve

$$p=\frac{1}{2}\,\frac{1}{(\alpha-1)!}\left[\frac{(\alpha-1)!}{2^{\alpha-1}}+\frac{\alpha!}{1!\,2^{\alpha}}+\ldots+\frac{(2\alpha-2)!}{(\alpha-1)!\,2^{2\alpha-2}}\right],$$

et cette quantité est égale à $\frac{1}{2}$.

La même méthode s'appliquerait encore si, au lieu de deux joueurs, on en considérait un plus grand nombre et si l'on accordait le gain de toutes les mises au premier qui aurait amené un nombre de points, donné à l'avance et différent pour chacun d'eux.

Problème II. — A *joue avec* B *aux conditions sui-*

vantes : on jette deux dés sur une table ; lorsqu'on amène un point au-dessous de 10, B *donne à* A *autant de francs que l'on a amené de points; dans le cas contraire,* A *donne à* B *une somme fixe, qu'il s'agit de déterminer, de manière à rendre le jeu équitable.*

A attend les sommes 2, 3, 4,..., 9; la probabilité d'obtenir le point 2 est $\frac{1}{36}$, la probabilité d'obtenir 3 est $\frac{2}{36}$, car il y a deux manières d'amener ce point, à savoir : as, deux et deux, as. La probabilité du point 4 est $\frac{3}{36}$, car on peut l'amener de trois manières : as, trois; deux, deux et trois, as; etc. : l'importance de la somme exposée par B est donc

$$2\frac{1}{36}+3\frac{2}{36}+4\frac{3}{36}+5\frac{4}{36}+6\frac{5}{35}+7\frac{6}{36}+8\frac{5}{36}+9\frac{4}{36},$$

ou bien $\frac{188}{36}$. Si nous désignons alors par x la somme que A doit payer à B, nous trouverons, pour l'importance de cette somme, la valeur

$$x\left(\frac{3}{36}+\frac{2}{36}+\frac{1}{36}\right) \quad \text{ou} \quad \frac{6}{36}x.$$

Si nous égalons les importances des sommes attendues par A et B, nous aurons

$$188=6x,$$

d'où nous tirons pour x la valeur $31^{fr},33$.

PROBLÈME II. — A *et* B *jouent à pile ou face aux conditions suivantes : si* A *amène pile au premier coup, il reçoit* 1 *franc et le jeu cesse; s'il l'amène au second coup seulement il reçoit* 2 *francs, etc.; s'il l'amène au* $n^{ième}$ *coup*

seulement, il reçoit 2^{n-1} francs; quelle doit être la mise de A?

Si l'on applique la règle de l'espérance mathématique à cette question, on trouve que la valeur espérée par A se compose des termes $\frac{1}{2}$ pour le premier coup, $\frac{1}{4}$ 2 pour le second, etc., $\frac{1}{2^n}2^{n-1}$ pour le $n^{ième}$; la mise de A devant être égale à cette somme est, comme on le voit, infinie.

On a voulu voir, dans la solution que nous venons de donner, un paradoxe et le problème précédent est devenu célèbre sous le nom de *problème de Saint-Pétersbourg*. Nicolas Bernoulli avait imaginé, pour expliquer ce prétendu paradoxe, la notion de l'*espérance morale*, sur laquelle nous ne nous arrêterons pas; mais, pour peu que l'on réfléchisse, on s'apercevra bien vite: 1° que, A pouvant, en définitive, gagner une somme infinie, il n'est pas étonnant que sa mise soit infinie; 2° que les arguments qui nous ont permis de définir avec précision l'importance d'une somme éventuelle ne peuvent plus s'appliquer ici. Et, en effet, pour pouvoir dire que $\frac{1}{2^n}2^{n-1}$ est le prix qu'il faudrait payer le droit de faire la $n^{ième}$ épreuve, il faut admettre que les joueurs puissent faire ensemble un très-grand nombre de parties, en mettant au jeu des sommes très-grandes sur des probabilités qui pourraient toujours être regardées comme très-petites, par rapport à l'inverse du nombre des parties jouées. Nous sommes alors en présence d'un cas où le théorème de Bernoulli tombe en défaut; c'est là, je crois, qu'il faut voir l'explication d'un fait qui répugne au sens commun et il faut en conclure que le jeu de Saint-Pétersbourg ne peut jamais être équitable.

Problème III. — *Un bijoutier possède un diamant brut dont la valeur est a francs; on vient lui annoncer que ce diamant est cassé en deux morceaux; un amateur, présent à ce moment, lui propose d'acheter les deux morceaux; on demande quel prix il devra les lui faire payer, avant d'avoir comparé leurs volumes?*

On sait que le prix du diamant varie comme le carré de son poids ou, si l'on veut, de son volume et, si le fait n'est pas très-exact, rien ne nous empêche de l'admettre dans le cas actuel. Prenons pour unité le volume total du diamant primitif et soient x et y les volumes des morceaux, en sorte que

$$x + y = 1. \tag{1}$$

Les prix de ces morceaux seront respectivement ax^2 et ay^2 et, par suite, le prix du diamant brisé sera

$$a(x^2 + y^2).$$

Or la probabilité que le volume du premier morceau sera compris entre x et $x + dx$ est dx; l'espérance mathématique du bijoutier dans ce cas est

$$a(x^2 + y^2)\,dx.$$

Si l'on fait la somme de toutes ces espérances pour les valeurs de x comprises entre 0 et 1, on aura le prix que l'on devra équitablement payer le diamant brisé, à savoir :

$$a\int_0^1 (x^2 + y^2)\,dx.$$

En remplaçant y par sa valeur tirée de la formule (1), cette

expression devient

$$a\int_0^1 (2x^2-2x+1)\,dx = \frac{2}{3}a.$$

Ainsi un diamant cassé en deux morceaux perd un tiers de sa valeur. En réalité, il devrait perdre davantage, à cause de l'irrégularité de la forme des fragments.

Supposons maintenant le diamant cassé en trois morceaux; en raisonnant comme tout à l'heure et en désignant par x, y, z les volumes de ces morceaux, on trouvera, pour leur prix,

$$a(x^2+y^2+z^2),$$

et l'on aura

$$(2)\qquad x+y+z=1.$$

L'espérance du bijoutier relative à des morceaux, dont les grosseurs seraient comprises entre x et $x+dx$, y et $y+dy$, sera

$$a(x^2+y^2+z^2)\,dx\,dy,$$

et l'importance du prix que peut espérer le bijoutier

$$a\iint (x^2+y^2+z^2)\,dx\,dy,$$

l'intégrale étant prise pour toutes les valeurs positives de x, y et z, satisfaisant à la formule (2); en d'autres termes, il faudra calculer l'intégrale

$$a\iint (2x^2+2y^2-2x-2y+2xy+1)\,dx\,dy,$$

pour les valeurs de x et y positives, telles que

$$x+y<1;$$

sa valeur est $\frac{a}{12}$.

Problème des partis.

Des joueurs se séparent quelquefois avant d'avoir terminé leur partie; le problème des *partis* a pour but de déterminer dans quel rapport ils doivent se partager l'enjeu. La règle de l'espérance mathématique donne immédiatement la solution de la question; en effet, supposons que des joueurs A, B, C,..., ayant la probabilité $p, q, r,\dots$ de gagner l'enjeu e, veuillent vendre leur droit de jouer à d'autres joueurs A', B', C',..., ces derniers devront payer à ceux dont ils acceptent la succession des sommes respectives pe, qe, re,... égales à leurs espérances mathématiques. Donc pe, qe, re,... sont les parts afférentes aux joueurs A, B, C,..., lorsqu'ils veulent se séparer avant d'avoir terminé leur partie. On conclut de là l'énoncé que voici :

Règle des partis. — *Lorsque des joueurs se séparent avant d'avoir terminé leur partie, ils doivent partager l'enjeu proportionnellement aux probabilités qu'ils auraient de gagner si la partie devait être achevée.*

Eclairons ceci par un exemple :

Problème. — *Deux joueurs,* A, B, *jouent en un certain nombre de points; ils veulent cesser leur partie au moment où il leur reste respectivement* α *et* β *points à marquer : soit* p *la probabilité que* A *a de marquer* 1 *point,* q *celle que* B *a de marquer* 1 *point; ces probabilités restant constantes pendant toute la durée du jeu, on demande dans quel rapport* B *et* A *doivent se partager l'enjeu.*

Supposons que l'on fasse continuer le jeu jusqu'à ce que A et B aient marqué ensemble $\alpha+\beta-1=s$ points. Il est

clair qu'alors l'un d'eux aura marqué les points qui lui restent à faire et que l'autre n'aura pas marqué les siens; il résulte de là que la probabilité que A gagnera et que B perdra ne sera pas altérée, en supposant que l'on continue le jeu jusqu'au $s^{ième}$ point. Or, si l'on considère le développement

$$(p+q)^s = p^s + \frac{s}{1}p^{s-1}q + \frac{s(s-1)}{1.2}p^{s-2}q^2 + \dots$$
$$+ \frac{s!}{\alpha!(\beta-1)!}p^\alpha q^{\beta-1} + \dots,$$

la somme des β premiers termes sera la probabilité que A marquera s points sur s, ou qu'il en marquera $s-1$ sur $s, \dots$; en d'autres termes, qu'il marquera au moins α points sur s, c'est-à-dire qu'il gagnera. La somme des termes suivants sera la probabilité que B gagnera. L'enjeu devra être partagé proportionnellement à ces probabilités.

On peut encore résoudre la question comme il suit : A peut gagner : 1° en marquant α points de suite; la probabilité de cet événement est p^α. 2° Il peut gagner en marquant $\alpha-1$ points dans les α premiers coups et 1 point au $\alpha+1^{ième}$; la probabilité de cet événement est

$$\frac{\alpha}{1}p^{\alpha-1}q \times p \quad \text{ou} \quad \frac{\alpha}{1}p^\alpha q.$$

3° Il peut gagner en marquant $\alpha-1$ points dans les $\alpha+1$ premiers coups et 1 point au $\alpha+2^{ième}$ coup; la probabilité de cet événement est

$$\frac{(\alpha+1)\alpha}{1.2}p^{\alpha-1}q^2 \times p \quad \text{ou} \quad \frac{\alpha(\alpha+1)}{1.2}p^\alpha q^2, \dots.$$

En appliquant le principe de la probabilité totale, on trouve,

pour la probabilité que A gagnera,

$$p^\alpha + \frac{\alpha}{1} p^\alpha q + \frac{\alpha(\alpha+1)}{1.2} p^\alpha q^2 + \ldots + \frac{\alpha(\alpha+1)\ldots(\alpha+\beta-2)}{1.2\ldots(\beta-1)} q^{\beta-1}.$$

Si l'on rapproche ce résultat de la solution précédente, on trouve, en égalant entre elles les deux formes que prend la probabilité que A gagnera, une identité remarquable due à Poisson; nous l'écrirons en changeant β en $\beta+1$, et nous aurons

$$p^s + \frac{s}{1} p^{s-1} q + \ldots + \frac{s!}{\alpha!\,\beta!} p^\alpha q^\beta$$
$$= p^\alpha \left[1 + \frac{\alpha}{1} q + \frac{\alpha(\alpha+1)}{1.2} q^2 + \ldots + \frac{\alpha(\alpha+1)\ldots(\alpha+\beta-1)}{1.2.3\ldots\beta} q^\beta \right].$$

Si au lieu de deux joueurs on en considérait un plus grand nombre, les conditions du jeu pourraient se compliquer selon qu'après un même coup un ou plusieurs joueurs marqueraient 1 point. Nous résoudrons seulement la question suivante :

A, B, C, ... *ont encore respectivement* $\alpha, \beta, \gamma, \ldots$ *points à faire pour gagner une partie déjà commencée; ils ont des probabilités respectives* $p, q, r, \ldots$ *de marquer* 1 *point à chaque coup, un seul d'entre eux marquant ce point, en sorte que* $p+q+r+\ldots=1$; *on demande comment ils doivent se partager l'enjeu.*

Il suffit, pour résoudre ce problème, de calculer la probabilité que chaque joueur aurait de gagner si la partie se continuait. A, par exemple, peut gagner : 1° en marquant α points de suite, ce dont la probabilité est p^α; 2° en marquant $\alpha-1$ points en α coups et 1 point au $\alpha+1^{\text{ième}}$, ce

dont la probabilité est

$$\sum \alpha p^{\alpha-1} q . q \quad \text{ou} \quad \sum \frac{\alpha!}{1.\nu!} p^{\alpha} q^{\nu},$$

le signe Σ s'étendant à toutes les valeurs q, r, s,..., et ν désignant l'unité; 3° en marquant $\alpha - 1$ points en $\alpha + 1$ coups et 1 point au $\alpha + 2^{ième}$ coup, ce dont la probabilité est

$$\sum \frac{\alpha!}{1.\nu!\,\omega!} p^{\alpha} q^{\nu} r^{\omega};$$

le signe Σ s'étend cette fois à toutes les valeurs q, r, s, ..., les exposants ν et ω pouvant être 0, 1 ou 2, mais leur somme faisant toujours 2. Ajoutons enfin qu'il faudrait exclure de dessous le signe Σ les termes en q^2 si $\beta = 2$, les termes en r^2 si $\gamma = 2$, etc., absolument comme sous le premier Σ que nous avons rencontré; il aurait fallu supprimer les termes en q si l'on avait eu $\beta = 1$, les termes en r si l'on avait eu $\gamma = 1$. Enfin il est clair aussi que, si l'on avait dû supprimer les termes en q sous le premier signe Σ, on devrait les supprimer aussi sous le second, ainsi que les termes en q^2; *a fortiori*, en poursuivant ce raisonnement et en appliquant le principe de la probabilité totale, on trouve, pour la probabilité que A gagnera,

$$p^{\alpha}\left(1 + \sum \frac{\alpha!}{\nu!} q^{\nu} + \sum \frac{\alpha!}{\nu!\,\omega!} q^{\nu} r^{\omega} + \ldots\right),$$

les signes Σ étant compris, comme nous l'avons expliqué tout à l'heure.

Le problème des partis peut se compliquer si la probabilité des joueurs varie à chaque coup; mais il se résout toujours par les mêmes principes. Par exemple, si deux individus, A, B, jouaient à rouge ou noire avec un jeu de

piquet, la probabilité de tirer rouge ou noire varierait avec la nature des cartes déjà sorties, c'est-à-dire avec la composition du paquet d'où l'on tire les cartes. Il est facile de voir, par exemple, que la probabilité que A aurait de gagner en pariant toujours pour rouge, quand il reste N cartes noires et R cartes rouges, serait la somme des probabilités :

1°
$$\frac{R}{N+R}\,\frac{R-1}{N+R-1}\cdots\frac{R-\alpha+1}{N+R-\alpha+1}$$

qu'il sortira α rouges de suite en α coups;

2°
$$\frac{\alpha N.R(R-1)\ldots(R-\alpha+2)}{(N+R)(N+R-1)\ldots(N+R-\alpha+1)}\,\frac{R-\alpha+1}{N+R-\alpha}$$

qu'il sortira $\alpha-1$ rouges dans les α premiers coups et 1 rouge au $\alpha+1^{ième}$, et ainsi de suite.

CHAPITRE IV.

SUR LES MÉTHODES DANS LES SCIENCES D'OBSERVATION.

Notions préliminaires.

Les erreurs que l'on commet forcément en opérant une mesure peuvent avoir leur origine dans l'imperfection des appareils dont on fait usage; mais ces erreurs sont alors ou constantes ou périodiques à fort peu de chose près, et l'on parvient souvent à les éliminer; on leur donne le nom d'*erreurs systématiques*.

Mais le défaut d'habileté de l'opérateur, l'imperfection de ses sens sont des causes d'erreurs dont l'influence se fait toujours sentir, et que le Calcul des probabilités seul peut apprendre à corriger. L'imperfection des instruments produit aussi des erreurs, très-faibles il est vrai, et qui n'ont rien de constant; d'autres causes, telles que l'influence des milieux où l'on opère, de la température, etc., peuvent, en agissant tantôt dans un sens, tantôt dans un autre, engendrer des erreurs qui ne semblent soumises à aucune loi régulière; on donne à ces erreurs qui n'ont rien de régulier le nom d'*erreurs accidentelles*.

Quoique au premier abord une erreur de mesure ne semble être soumise à aucune loi, il est bien évident pour tout le monde que les grosses erreurs se renouvelleront moins souvent que les petites, en sorte que la probabilité d'une erreur est d'autant plus forte que cette erreur est plus petite.

Si l'on considère l'erreur que l'on a pu commettre dans une mesure, il est évident que la probabilité que cette erreur est comprise entre o et x sera d'autant plus forte que x sera plus grand en valeur absolue, en sorte que la probabilité en question est une véritable fonction de x, inconnue, il est vrai, et variable avec les instruments et les procédés de mesure. Désignons alors par $F(x)$ la probabilité que l'erreur est comprise entre o et x.

Si l'on veut trouver la probabilité y que l'erreur est comprise entre x et $x+dx$, on observera que, en vertu du principe de la probabilité totale, la probabilité $F(x+dx)$ que l'erreur est comprise entre o et $x+dx$ est égale à la probabilité $F(x)$ que l'erreur est comprise entre o et x plus la probabilité y que l'erreur est comprise entre x et $x+dx$ (les cas favorables à l'existence de l'erreur comprise entre o et $x+dx$ se décomposant en deux séries s'excluant mutuellement, à savoir : les cas où l'erreur est comprise entre o et x, et les cas où elle est comprise entre x et $x+dx$) ; on a donc

$$F(x+dx) = F(x) + y,$$

d'où l'on conclut

$$y = F'(x)dx.$$

$F'(x)$ est ce que nous appellerons la *facilité de l'erreur* x.

Ainsi la *facilité de l'erreur* est une fonction $\varphi(x)$, telle que $\varphi(x)dx$ représente la probabilité que l'erreur est comprise entre x et $x+dx$.

Supposons que l'on fasse un très-grand nombre s d'observations pour déterminer une quantité O. Si l'on désigne par n_x le nombre de fois que l'erreur sera comprise entre x et $x+dx$, on aura à peu près, en vertu du théorème de

Bernoulli,

$$(1)\qquad \varphi(x)\,dx = \frac{n_x}{s}, \quad n_x = s\varphi(x)dx;$$

la somme des erreurs comprises entre x et $x+dx$ sera $n_x x$, c'est-à-dire $sx\varphi(x)\,dx$; la somme des carrés, des cubes, etc., de ces erreurs sera $sx^2\varphi(x)dx$, $sx^3\varphi(x)dx$.... Si l'on considère l'intégrale

$$M_0 = \int_{-\infty}^{+\infty} \varphi(x)\,dx,$$

cette intégrale sera égale à la somme des valeurs de n_x divisé par s, c'est-à-dire à $s:s$ ou à l'unité. On peut le prouver d'une autre manière, en observant que M_0 est la somme des probabilités que l'erreur est comprise dans chacun des intervalles compris entre $-\infty$ et $+\infty$, où la probabilité que cette erreur est comprise entre $-\infty$ et $+\infty$; or cette probabilité est une certitude : sa valeur est donc l'unité. L'intégrale

$$M_1 = \int_{-\infty}^{+\infty} x\varphi(x)dx$$

est la somme des valeurs que peut prendre xn_x divisé par s; cette quantité est donc égale à la moyenne des erreurs commises sur s observations.

$$M_2 = \int_{-\infty}^{+\infty} x^2\varphi(x)dx$$

sera la moyenne des carrés des erreurs, et ainsi de suite. Lorsque les erreurs positives ont la même facilité que les erreurs négatives de même valeur absolue, ce qui a ordinairement lieu, on a

$$M_1 = 0, \quad M_3 = 0, \ldots .$$

Lorsque le nombre des observations n'est pas infini, on conserve aux intégrales M_1, M_2,... les noms de *moyenne des erreurs, moyenne des carrés des erreurs*, etc.

L'*erreur moyenne*, qu'il ne faut pas confondre avec la *moyenne des erreurs*, est la quantité m définie par l'équation

$$\int_{-\infty}^{+\infty} \varphi(x) x^2 dx = m^2 \quad \text{ou} \quad m = \sqrt{M_2}.$$

La précision des observations sera d'autant plus grande, évidemment, que la somme des carrés des erreurs sera plus petite; on peut donc prendre $\frac{1}{m}$ pour la *mesure de la précision;* $\frac{1}{m^2}$ est ce que l'on appelle *poids* des observations.

L'*erreur probable* est l'erreur ε, déterminée par la relation

$$\frac{1}{2} = \int_{-\varepsilon}^{+\varepsilon} \varphi(x) dx;$$

c'est, comme on le voit, une quantité telle qu'il y ait 1 à parier contre 1 que l'erreur sera inférieure à cette quantité ε en valeur absolue : en effet l'intégrale précédente est la somme des probabilités que l'erreur sera comprise dans l'un des intervalles dx compris entre $-\varepsilon$ et $+\varepsilon$.

Recherche de la moyenne des erreurs.

Soit $\varphi(x)$ la facilité de l'erreur x dans une série de s observations $O_1, O_2, \ldots, O_s$, le produit

$$(1) \qquad \varphi(x_1) dx_1 \varphi(x_2) dx_2 \ldots \varphi(x_s) dx_s$$

exprimera la probabilité que l'erreur commise sur O_1 est

comprise entre x_1 et $x_1 + dx_1$, que l'erreur commise sur x_2 est comprise entre x_2 et $x_2 + dx_2, \ldots$ Si dans ce produit on fait varier $x_1, x_2, \ldots, x_s$ de telle sorte que leur somme fasse μ, la somme de tous les résultats ainsi obtenus sera la probabilité que la somme des erreurs commises sur O_1, $O_2, \ldots, O_s$ est μ, ou diffère de μ d'une quantité moindre que $dx_1 + dx_2 + \ldots + dx_s$. Si enfin on fait varier μ entre $\alpha - l$ et $\alpha + l$, ou, ce qui revient au même, si l'on fait varier $x_1, x_2, \ldots, x_s$ de telle sorte que l'on ait

$$(2) \qquad \alpha - l < x_1 + x_2 + \ldots + x_s < \alpha + l,$$

la somme des produits tels que (1) sera la probabilité que la somme des erreurs commises sur $O_1, O_2, \ldots, O_s$ est comprise entre $\alpha - l$ et $\alpha + l$. En désignant cette probabilité par P, on aura donc

$$(3) \qquad P = \int\int\int \ldots \varphi(x_1)\varphi(x_2)\ldots dx_1\, dx_2 \ldots dx_s,$$

l'intégrale multiple étant prise de manière à satisfaire à la condition (2). L'intégrale multiple (3), prise entre des limites variables, peut être ramenée, par une méthode employée par Dirichlet dans d'autres circonstances, à une intégrale prise entre des limites constantes; il suffit pour cela de multiplier l'élément placé sous le signe d'intégration par un facteur égal à 1 quand Σx est compris entre $\alpha - l$ et $\alpha + l$, et égal à zéro en dehors de ces limites.

Il est clair que l'on pourra alors intégrer entre les limites $-\infty$ et $+\infty$, les éléments ajoutés étant nuls. Le facteur en question sera fourni par la formule de Fourier (p. 31)

$$\psi(t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \psi(\tau) e^{2\pi u (t-\tau)\sqrt{-1}} d\tau\, du.$$

Si dans cette formule on fait $\psi(t) = 1$ ou 0, selon que t est ou n'est pas compris entre $\alpha - l$ et $\alpha + l$, on a

$$\psi(t) = \int_{-\infty}^{+\infty} \int_{\alpha-l}^{\alpha+l} e^{2\pi u(t-\tau)\sqrt{-1}}\, d\tau\, du,$$

ou, en effectuant la première intégration,

$$\psi(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} e^{2u\pi(t-\alpha)\sqrt{-1}} \frac{\sin 2u\pi l}{u}\, du.$$

Si l'on suppose $t = x_1 + x_2 + \ldots + x_s$, la fonction $\psi(t)$ sera précisément le facteur dont nous avons besoin, et la formule (3) pourra s'écrire

$$\begin{aligned} \mathrm{P} = \frac{1}{\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} & \ldots \varphi(x_1)\varphi(x_2)\ldots \\ & \times \varphi(x_s) e^{2u\pi(x_1+x_2+\ldots+x_s-\alpha)\sqrt{-1}} \\ & \times \frac{\sin 2u\pi l}{u}\, du\, dx_1\, dx_2 \ldots dx_s; \end{aligned}$$

on peut, sans inconvénient, remplacer $2u\pi$ par u en changeant de variable, et l'on a alors

$$\begin{aligned} \mathrm{P} = \frac{1}{\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} & \ldots \varphi(x_1) e^{ux_1\sqrt{-1}}\, dx_1 \\ & \times \varphi(x_2) e^{ux_2\sqrt{-1}}\, dx_2 \ldots e^{-u\alpha\sqrt{-1}} \frac{\sin ul}{u}\, du. \end{aligned}$$

Il est sous-entendu que $\varphi(x)$ est nul pour $x = \infty$, les erreurs infinies étant impossibles : on peut donc intervertir l'ordre des intégrations et écrire

$$(4) \qquad \mathrm{P} = \frac{1}{\pi} \int_{-\infty}^{+\infty} \mathrm{T}^s e^{-\alpha u\sqrt{-1}} \frac{\sin ul}{u}\, du,$$

T désignant, pour abréger, l'intégrale

$$(5)\qquad T = \int_{-\infty}^{+\infty} \varphi(x) e^{ux\sqrt{-1}} dx.$$

Laplace et Poisson arrivent au même résultat, mais par une voie toute différente. Nous pouvons maintenant faire diverses hypothèses sur la nature de la fonction φ :

Supposons $\varphi(x)$ constant entre les limites $-\frac{1}{2}$ et $+\frac{1}{2}$, toutes les erreurs sont alors aussi probables les unes que les autres; ce cas se trouve réalisé dans la pratique, lorsque l'on fait dans les Compagnies financières une répartition de bénéfices entre les actionnaires, et que cette répartition se fait par *nombres ronds* de francs. L'erreur est sur chaque résultat moindre que $\frac{1}{2}$ franc, et toutes les erreurs comprises entre $-\frac{1}{2}$ et $+\frac{1}{2}$ sont également probables. Dans ce cas, la formule (5) donne

$$T = k\int_{-\frac{1}{2}}^{+\frac{1}{2}} e^{ux\sqrt{-1}} dx = 2\frac{k}{u}\sin\frac{1}{2}u;$$

la constante k se détermine en observant que

$$\int_{-\infty}^{+\infty} \varphi(x) dx = 1 \quad \text{ou} \quad \int_{-\frac{1}{2}}^{+\frac{1}{2}} k\,dx = 1,$$

d'où l'on tire $k = 1$.

La formule (4) devient alors

$$P = \frac{1}{\pi}\int_{-\infty}^{+\infty} \left(2\frac{\sin\frac{1}{2}u}{u}\right)^{s} e^{-au\sqrt{-1}} \frac{\sin ul}{u} du;$$

en prenant $\alpha = 0$, on aura la probabilité que

$$x_1 + x_2 + \ldots + x_s$$

reste inférieur à l en valeur absolue ou

$$(6) \qquad P = \frac{1}{\pi} \int_{-\infty}^{+\infty} \left(2 \frac{\sin \frac{1}{2} u}{u} \right)^s \frac{\sin ul}{u} du.$$

On pourrait, si l'on y tenait, obtenir le développement de P en série : il suffirait, pour cela, de partager l'intégrale en plusieurs autres, dont les limites différeraient de 2π et que l'on évaluerait en les développant suivant les puissances de u, de $2\pi - u$, $4\pi - u, \ldots$; mais, dans les applications, il est plus avantageux de se contenter d'une solution dont l'approximation ne peut pas être poussée aussi loin que l'on veut, à la vérité, mais qui donne immédiatement une idée de l'erreur admissible. Observons que l'intégrale (6) n'a de valeur sensible que pour les très-petites valeurs de u lorsque s est très-grand; on a alors

$$\frac{2 \sin \frac{1}{2} u}{u} = \frac{2}{u} \left(\frac{1}{2} u - \frac{1}{8} \frac{u^3}{1.2.3} + \ldots \right),$$

ou, à fort peu de chose près,

$$\frac{2 \sin \frac{1}{2} u}{u} = 1 - \frac{u^2}{24} = e^{-\frac{u^2}{24}},$$

et, par suite, l'intégrale (6) devient

$$P = \frac{1}{\pi} \int_{-\infty}^{+\infty} e^{-\frac{su^2}{24}} \frac{\sin ul}{u} du.$$

On démontre dans les Cours d'Analyse que

$$\int_{-\infty}^{+\infty} e^{-au^2} \cos lu\, du = \sqrt{\frac{\pi}{a}}\, e^{-\frac{l^2}{4a}};$$

en intégrant par rapport à l, cette formule devient

$$\int_{-\infty}^{+\infty} e^{-au^2} \frac{\sin lu}{u}\, du = \sqrt{\frac{\pi}{a}} \int_0^l e^{-\frac{l^2}{4a}}\, dl.$$

En faisant $a = \frac{s}{24}$, on a

$$\int_{-\infty}^{+\infty} e^{-\frac{s}{24}u^2} \frac{\sin lu}{u}\, du = \sqrt{\frac{24\pi}{s}} \int_0^l e^{-\frac{6l^2}{s}}\, dl$$

et, par suite,

$$\mathrm{P} = \sqrt{\frac{24}{s\pi}} \int_0^l e^{-\frac{6l^2}{s}}\, dl.$$

Si l'on pose, en général,

$$\Theta(\gamma) = \frac{2}{\sqrt{\pi}} \int_0^\gamma e^{-\gamma^2} d\gamma,$$

la formule précédente devient

$$\mathrm{P} = \Theta\left(l\sqrt{\frac{6}{s}}\right).$$

Posons (*voir* p. 34)

$$l\sqrt{\frac{6}{s}} = 2{,}4 \quad \text{d'où} \quad \mathrm{P} = 0{,}999,$$

et il y aura mille à parier contre un que l'erreur totale sera comprise entre $-l$ et $+l$. Supposons, par exemple, $s = 600$, on pourra parier mille contre un que l'erreur totale est com-

prise entre — 24 et + 24. En résumé, nous pouvons énoncer la règle suivante :

RÈGLE. — *Pour faire la preuve d'une règle de société ou d'un inventaire, ou plus généralement pour faire la preuve d'une addition dont on connaît le résultat exact, mais dans laquelle les parties à ajouter ont été remplacées par des* NOMBRES RONDS, *formez le radical* $\sqrt{\frac{s}{6}}$, *s désignant le nombre des parties, multipliez-le par* 2,4, *et il y aura plus de mille à parier contre un que, si les parties sont convenablement approchées, et si l'addition est bien faite, l'erreur, c'est-à-dire la différence entre la somme exacte et la somme approchée sera inférieure à* $2,4\sqrt{\frac{s}{6}}$.

Revenons au cas général : l'intégrale (5) peut être développée en série, et l'on a

$$\mathrm{T} = \int_{-\infty}^{+\infty} \varphi(x)\left(1 + \frac{ux\sqrt{-1}}{1} - \frac{u^2x^2}{1.2} + \ldots\right)dx,$$

et, en faisant usage d'une notation employée plus haut,

$$\mathrm{T} = 1 + \frac{\mathrm{M}_1 u\sqrt{-1}}{1} - \frac{\mathrm{M}_2 u^2}{1.2} - \frac{\mathrm{M}_3 u^3\sqrt{-1}}{1.2.3} + \ldots.$$

Généralement $\varphi(x)$ décroît rapidement à mesure que x croît, les grosses erreurs étant moins à craindre que les petites, et les intégrales $\mathrm{M}_1, \mathrm{M}_2, \mathrm{M}_3, \ldots$ deviennent de plus en plus petites quand leur indice augmente : on aura donc, à peu de chose près et pour de petites valeurs de u,

$$\mathrm{T} = 1 + \mathrm{M}_1 u\sqrt{-1} - \frac{\mathrm{M}_2 u^2}{2},$$

ou, à très-peu près,

$$\log T = M_1 u \sqrt{-1} - \frac{M_2 u^2}{2} + \frac{M_1^2 u^2}{2},$$

d'où l'on conclut

$$T = e^{M_1 u \sqrt{-1}} e^{-u^2 \frac{M_2 - M_1^2}{2}};$$

la formule (4), dans laquelle l'intégrale n'a de valeurs sensibles que pour de petites valeurs de u (et surtout si s est très-grand, le module de T étant moindre que 1 dès que u est différent de zéro), devient alors

$$P = \frac{1}{\pi} \int_{-\infty}^{+\infty} e^{(s M_1 - \alpha) u \sqrt{-1}} e^{-u^2 \frac{M_2 - M_1^2}{2} s} \frac{\sin ul}{u} du.$$

Supposons que l'on prenne α égal à $s M_1$, on aura

$$P = \frac{1}{\pi} \int_{-\infty}^{+\infty} e^{-u^2 a} \frac{\sin ul}{u} du,$$

en désignant par a la quantité $s \frac{M_2 - M_1^2}{2}$. La valeur de P est donc (p. 34)

$$P = \sqrt{\frac{1}{a\pi}} \int_0^l e^{-\frac{l^2}{4a}} dl, \quad \text{ou} \quad P = \frac{2}{\sqrt{\pi}} \int_0^\gamma e^{-\gamma^2} d\gamma,$$

formule où l'on a posé

$$\gamma = \frac{l}{\sqrt{2s(M_2 - M_1^2)}}.$$

On peut observer que M_1 est la *moyenne des erreurs* (de toutes les erreurs que l'on peut commettre sur un nombre infini d'observations) : on peut donc dire que la moyenne μ des erreurs commises sur s observations diffère

très-probablement de la moyenne M_1 d'une quantité inférieure à $\frac{l}{s}$, l étant donné par la formule

$$\frac{l}{\sqrt{2s(M_2 - M_1^2)}} = \gamma.$$

Si l'on prend $\gamma = 2,4$, il y aura mille à parier contre un que μ diffère de M_1 de moins de $\frac{l}{s}$.

Il est essentiel d'observer que l'analyse précédente suppose les quantités $M_0, M_1, M_2, \ldots$ rapidement convergentes vers zéro, et c'est ce qui a lieu effectivement pour les erreurs d'observations.

Recherche de la moyenne des carrés des erreurs.

Conservons les mêmes notations que dans le paragraphe précédent, et proposons-nous de trouver la probabilité pour que la somme des carrés des erreurs soit comprise entre des limites données $\alpha - l$ et $\alpha + l$.

On verra, en suivant la même analyse que tout à l'heure, que la probabilité cherchée est donnée par la formule

$$P = \int\int\int \ldots \varphi(x_1)\,dx_1\, \varphi(x_2)\,dx_2 \ldots \varphi(x_s)\,dx_s;$$

mais cette fois les intégrales devront être prises de manière à satisfaire à la relation

$$\alpha - l < x_1^2 + x_2^2 + \ldots + x_s^2 < \alpha + l.$$

Pour pouvoir effectuer l'intégration, nous introduirons encore un facteur nul quand l'inégalité précédente n'est pas satisfaite, et égal à 1 quand elle l'est : ce facteur est toujours

donné par la formule de Fourier, et sa valeur est

$$\frac{1}{\pi}\int_{-\infty}^{+\infty} e^{2u\pi(t-a)\sqrt{-1}}\,\frac{\sin 2u\pi l}{u}\,du,$$

t représentant, pour abréger, $x_1^2 + x_2^2 + \ldots + x_s^2$: on a donc

$$\mathrm{P} = \frac{1}{\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\ldots\varphi(x_1)e^{2u\pi x_1^2\sqrt{-1}}dx_1\,\varphi(x_2)e^{2u\pi x_2^2\sqrt{-1}}dx_2\ldots$$

$$\times\, e^{2u\pi a\sqrt{-1}}\,\frac{\sin 2u\pi l}{u}\,du,$$

ou bien

$$\mathrm{P} = \frac{1}{\pi}\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty}\varphi(x)e^{2u\pi x^2\sqrt{-1}}dx\right]^s e^{-2u\pi a\sqrt{-1}}\,\frac{\sin 2u\pi l}{u}\,du,$$

ou bien encore, en changeant $2u\pi$ en u,

$$(1)\quad \mathrm{P} = \frac{1}{\pi}\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty}\varphi(x)e^{ux^2\sqrt{-1}}dx\right]^s e^{-ua\sqrt{-1}}\,\frac{\sin ul}{u}\,du.$$

Or on a

$$\int_{-\infty}^{+\infty}\varphi(x)e^{ux^2\sqrt{-1}}dx$$

$$= \int_{-\infty}^{+\infty}\varphi(x)\left[1 + \frac{ux^2\sqrt{-1}}{1} - \frac{u^2x^4}{1.2} + \ldots\right]dx$$

$$= 1 + \mathrm{M}_2 u\sqrt{-1} - \mathrm{M}_4\frac{u^2}{1.2} + \ldots,$$

ou à peu près

$$e^{\mathrm{M}_2 u\sqrt{-1}}\,e^{-\frac{1}{2}(\mathrm{M}_4 - \mathrm{M}_2^2)u^2};$$

par suite, la formule (1) devient

$$P = \frac{1}{\pi} \int_{-\infty}^{+\infty} e^{(s M_2 - \alpha) u \sqrt{-1}} e^{-\frac{s}{2}(M_4 - M_2^2) u^2} \frac{\sin ul}{u} du.$$

Si l'on prend $\alpha = M_2 s$, on aura, en suivant l'analyse développée au paragraphe précédent,

$$P = \frac{2}{\sqrt{\pi}} \int_0^{\gamma} e^{-\gamma^2} d\gamma \quad \text{avec} \quad \gamma = \frac{l}{\sqrt{2s(M_4 - M_2^2)}},$$

et il est facile d'en conclure que très-probablement la moyenne des carrés des erreurs d'un grand nombre d'observations s'écartera fort peu de M_2. Cette remarque nous sera utile un peu plus loin.

On voit sans peine comment on pourrait trouver la probabilité que la somme des cubes des erreurs fût comprise entre des limites données. Disons, pour terminer, comment on pourrait trouver la probabilité que la somme des valeurs absolues des erreurs fût comprise entre $\alpha - l$ et $\alpha + l$. Il faudrait évidemment évaluer l'intégrale multiple

$$\iint \ldots \varphi(x_1) dx_1 \, \varphi(x_2) dx_2 \ldots,$$

de telle sorte que l'on eût

$$\alpha - l < \sqrt{x_1^2} + \sqrt{x_2^2} + \ldots + \sqrt{x_s^2} < \alpha + l,$$

et la question ainsi présentée se traitera comme les précédentes.

Sur la moyenne d'un grand nombre d'observations.

Conservons encore les notations des paragraphes qui précèdent. Nous avons vu (p. 44) que la probabilité que la

somme des erreurs commises sur s observations est comprise entre $sM_1 - l$ et $sM_1 + l$ était donnée par la formule

$$P = \Theta\left[\frac{l}{\sqrt{2s(M_2 - M_1^2)}}\right];$$

c'est la probabilité que la moyenne des erreurs est comprise entre $M_1 - \frac{l}{s}$ et $M_1 + \frac{l}{s}$. Si nous posons $\frac{l}{s} = \lambda$, nous aurons

$$P = \Theta\left[\frac{\lambda\sqrt{s}}{\sqrt{2(M_2 - M_1^2)}}\right],$$

et P représentera la probabilité que la moyenne des s erreurs est comprise entre $M_1 - \lambda$ et $M_1 + \lambda$. Or, en appelant $m_1, m_2, \ldots, m_s$ les mesures données par l'observation de la quantité ω que l'on cherche, on a

$$x_1 = \omega - m_1, \quad x_2 = \omega - m_2, \ldots, \quad x_s = \omega - m_s,$$

d'où l'on conclut, en ajoutant et divisant par s,

$$\text{moy}.x = \omega - \text{moy}.m.$$

Ainsi la différence entre ω et la moyenne des observations est la moyenne des erreurs; on peut donc dire que :

THÉORÈME. — *Lorsque l'on cherche à mesurer une quantité* ω, *la différence entre cette quantité et la moyenne des observations est une quantité qui a la probabilité*

$$\Theta\left[\frac{\lambda\sqrt{s}}{\sqrt{2(M_2 - M_1^2)}}\right]$$

de se trouver comprise entre $M_1 - \lambda$ et $M_1 + \lambda$.

Lorsque les erreurs systématiques ont été éliminées, les

erreurs par excès et les erreurs par défaut deviennent aussi probables les unes que les autres; alors on a

$$M_1 = \int_{-\infty}^{+\infty} x\varphi(x)\,dx = 0,$$

et l'on a la probabilité

$$P = \Theta\left(\frac{\lambda\sqrt{s}}{\sqrt{2M_2}}\right)$$

que ω diffère de la moyenne des observations d'une quantité inférieure à λ. Si l'on se donne P très-voisin de l'unité, on voit que le produit $\lambda\sqrt{s}$ devra conserver une valeur constante liée à la valeur donnée de P, et, par suite, si s est très-grand, λ pourra être très-petit, ce qui tendrait à prouver que la moyenne d'un très-grand nombre d'observations approche constamment du résultat exact et avec une rapidité dont la mesure est la racine carrée du nombre des observations. Cette conclusion, il ne faut pas l'oublier, ne s'applique qu'au cas où les erreurs systématiques sont éliminées.

Pour calculer P, il faut, comme on le voit, connaître M_2, et il semble que la connaissance de la fonction φ soit indispensable; toutefois, si l'on se rappelle que la moyenne des carrés des erreurs s'écarte fort peu de M_2 (*voir* p. 152), on pourra, pour le calcul de P, remplacer M_2 par la moyenne des carrés des erreurs calculées *a posteriori*, c'est-à-dire que, m_i désignant la mesure fournie par l'observation O_i, on prendra pour valeur de M_1 la quantité $\frac{\Sigma m}{s} = \omega$, qui en diffère très-peu; l'erreur commise sur l'observation O_i sera sensiblement égale à $\omega - m_i$, et par suite la valeur approchée de M_2 sera $\frac{\Sigma(\omega - m_i)^2}{s}$.

Sur la détermination de la fonction φ.

Supposons que l'on ait éliminé avec soin les erreurs systématiques; nous avons vu que la probabilité que la valeur exacte ω diffère de la moyenne $\frac{m_1 + m_2 + \ldots + m_s}{s} = \mu$ des observations d'une quantité inférieure à λ était

$$\Theta\left(\lambda\sqrt{\frac{s}{2M_2}}\right).$$

Cette quantité peut être prise aussi voisine de l'unité que l'on veut en prenant s assez grand; il en résulte que, quand s est très-grand, la valeur la plus probable de ω est μ; mais ceci n'est vrai : 1° que si s est très-grand; 2° que si les erreurs systématiques ont été éliminées, ou, ce qui revient au même, que si $M_1 = 0$.

Gauss a démontré, dans la *Theoria motus corporum cœlestium,* que si l'on admettait *a priori* que la moyenne μ était la valeur la plus probable de ω, la fonction φ était déterminée à une constante près. Voici l'analyse de Gauss un peu simplifiée :

La probabilité du concours des erreurs $x_1 = \omega - m_1$, $x_2 = \omega - m_2, \ldots, x_s = \omega - m_s$ dans le courant de s observations est

$$\varphi(x_1)\ldots\varphi(x_s)\,dx_1\,dx_2\ldots dx_s,$$

ou, si l'on veut,

$$\varphi(\omega - m_1)\,\varphi(\omega - m_2)\ldots dx_1\ldots dx_s.$$

En vertu du théorème de Bayes, si l'on considère ω comme la *cause* des erreurs $x_1, x_2, \ldots$, la probabilité de

cette cause, quand on aura observé $m_1, m_2, \ldots$, sera

$$\frac{\varphi(x_1)\,\varphi(x_2)\ldots\varphi(x_s)\,dx_1\,dx_2\ldots dx_s}{\int\int\int_{-\infty}^{+\infty}\ldots\varphi(x_1)\,\varphi(x_2)\ldots dx_1\,dx_2\ldots dx_s}.$$

Le dénominateur est le produit de s facteurs égaux à l'unité et, par conséquent, la probabilité de la valeur ω de la quantité à mesurer est

$$\varphi(x_1)\ldots\varphi(x_s)\,dx_1\ldots dx_s,$$

ou, pour parler plus correctement, l'expression précédente est la probabilité que la quantité à mesurer est comprise entre ω et $\omega + d\omega$. La valeur la plus probable de ω est celle qui rendra

$$\varphi(x_1)\ldots\varphi(x_s) \quad \text{ou} \quad \varphi(\omega - m_1)\,\varphi(\omega - m_2)\ldots\varphi(\omega - m_s)$$

maximum, c'est-à-dire qui annulera la dérivée de cette quantité, ou

$$\frac{\varphi'(\omega - m_1)}{\varphi(\omega - m_1)} + \frac{\varphi'(\omega - m_2)}{\varphi(\omega - m_2)} + \ldots + \frac{\varphi'(\omega - m_s)}{\varphi(\omega - m_s)}.$$

Nous connaissons cette valeur, qui est, par hypothèse, la moyenne μ; on a donc une propriété de la fonction φ, qui est donnée par la relation

$$\sum \frac{\varphi'(\mu - m_i)}{\varphi(\mu - m_i)} = 0,$$

qui doit avoir lieu identiquement, c'est-à-dire quels que soient $m_1, m_2, \ldots$. Si l'on pose

$$\frac{\varphi'(x)}{\varphi(x)} = F(x),$$

on a

$$\Sigma F(\mu - m_i) = 0.$$

En différentiant, par rapport à $m_1, m_2, \ldots, m_s$, on trouve

$$\sum F'(\mu - m_i)\frac{d\mu}{dm_1} - F'(\mu - m_1) = 0,$$
$$\sum F'(\mu - m_i)\frac{d\mu}{dm_2} - F'(\mu - m_2) = 0,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots ;$$

mais, comme $\frac{d\mu}{dm_1} = \frac{d\mu}{dm_2} = \ldots = \frac{1}{s}$, on en conclut que

$$F'(\mu - m_1) = F'(\mu - m_2)\ldots,$$

c'est-à-dire que la fonction $F'(x)$ conserve la même valeur, quel que soit x : on peut donc poser

$$F'(x) = 2h,$$

et, par suite,

$$F(x) = 2hx + h',$$

h et h' désignant deux constantes. Remplaçant $F(x)$ par sa valeur, on a

$$\frac{\varphi'(x)}{\varphi(x)} = 2hx + h',$$

et, par suite, en intégrant,

$$\log\varphi(x) = hx^2 + h'x + h'',$$

d'où l'on tire

$$\varphi(x) = e^{hx^2 + h'x + h''}.$$

Si l'on fait $x \pm \infty$, $\varphi(x)$ doit s'annuler, ce qui exige : 1° que h soit négatif ; 2° que h' soit nul. En remplaçant alors h par $-h^2$ et h'' par $\log k$, on aura

$$\varphi(x) = ke^{-h^2x^2}.$$

On peut encore déterminer k en observant que l'intégrale de $\varphi(x)dx$ est égale à 1, entre les limites $-\infty$ et $+\infty$:

on a donc

$$\int_{-\infty}^{+\infty} k e^{-h^2 x^2} = 1,$$

ce qui peut s'écrire

$$\frac{k\sqrt{\pi}}{h} = 1, \quad \text{d'où} \quad k = \frac{h}{\sqrt{\pi}};$$

donc enfin

$$\varphi(x) = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2}.$$

La constante h ne peut pas être déterminée et, en effet, c'est d'elle que dépendra la précision des observations.

Si l'on calcule l'erreur moyenne $\sqrt{M_2}$, on trouve

$$M_2 = \int_{-\infty}^{+\infty} \frac{h x^2}{\sqrt{\pi}} e^{-h^2 x^2} dx = \frac{1}{2h^2}:$$

on a donc

$$h = \frac{1}{\sqrt{2 M_2}}.$$

C'est en partant de cette hypothèse sur la forme de la fonction φ, hypothèse gratuite (car *a priori* on pourrait choisir comme résultat le plus probable la moyenne géométrique avec autant de raison que la moyenne arithmétique) que Gauss est parvenu à justifier pour la première fois (1809, *Theoria motus, etc.*) la célèbre méthode des moindres carrés, imaginée en 1805 par Legendre et que nous allons exposer tout à l'heure.

Méthode des moindres carrés.

Dans un grand nombre de questions de Physique ou d'Astronomie, on est conduit à déterminer certaines quantités

à l'aide d'équations dont les coefficients sont donnés par l'expérience. Au lieu de faire le nombre d'expériences strictement nécessaire à la détermination des inconnues, en d'autres termes, au lieu de faire un nombre d'expériences égal au nombre des inconnues, afin d'avoir autant d'équations à résoudre que d'inconnues à déterminer, on multiplie en général le nombre des expériences autant que possible. On a alors des équations en nombre bien supérieur à celui des inconnues, mais qui ne sont pas rigoureusement compatibles, puisque les mesures ne comportent pas une précision absolue; et, de la discussion de ces équations, on conçoit qu'il puisse résulter plus d'exactitude que si l'on prenait seulement celles qui sont rigoureusement suffisantes à la détermination des inconnues. Nous allons maintenant montrer comment on peut obtenir cet accroissement de précision, en supposant d'abord que l'on ait affaire à des équations du premier degré.

Supposons donc qu'il s'agisse de déterminer n inconnues x, y, z,..., au moyen d'équations approximatives de la forme

$$(1) \quad \begin{cases} a_1 x + b_1 y + c_1 z + \ldots = k_1, \\ a_2 x + b_2 y + c_2 z + \ldots = k_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_s x + b_s y + c_s z + \ldots = k_s. \end{cases}$$

Ces équations, fournies par l'expérience, sont en plus grand nombre que les inconnues, et, si les observations qui les ont fournies étaient faites avec toute la perfection voulue, elles seraient compatibles; en réalité, elles ne le sont pas et il s'agit alors d'en déduire les valeurs les plus probables des inconnues.

Pour calculer x, nous multiplierons chacune de ces équations par un facteur particulier et nous ferons en sorte

qu'en les ajoutant le coefficient de x soit 1 et celui des autres inconnues zéro. Nous trouverons alors, en désignant par $\lambda_1, \lambda_2, \ldots, \lambda_s$ les facteurs en question,

$$(2) \qquad x = \Sigma\lambda k,$$

formule à laquelle nous devrons joindre les suivantes :

$$(3) \qquad \Sigma a\lambda = 1, \quad \Sigma b\lambda = 0, \quad \Sigma c\lambda = 0, \ldots.$$

Désignons par $\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots$ les erreurs qui affectent les quantités $k_1, k_2, \ldots$, et par $\xi, \eta, \zeta, \ldots$ les erreurs qui affectent les inconnues $x, y, z, \ldots$ données par la formule (2) et ses analogues; en sorte que, si l'on mettait pour $x, y, z, \ldots$ leurs valeurs exactes $x+\xi$, $y+\eta, \ldots$ dans (1), les seconds membres devraient être remplacés par $k_1+\varepsilon_1$, $k_2+\varepsilon_2, \ldots, k_s+\varepsilon_s$; on aura évidemment, de la même façon qu'on a obtenu la formule (2),

$$x + \xi = \Sigma\lambda(k + \varepsilon);$$

de cette formule et de la formule (2) on tire

$$(4) \qquad \xi = \Sigma\lambda\varepsilon.$$

Cherchons la probabilité que l'erreur ξ sera comprise entre des limites données $\alpha - l$ et $\alpha + l$. Soient $\varphi_1(\varepsilon_1)$ la facilité de l'erreur ε_1, $\varphi_2(\varepsilon_2)$ la facilité de l'erreur ε_2, etc., la probabilité du concours de ces erreurs sera

$$\varphi_1(\varepsilon_1)\varphi_2(\varepsilon_2)\ldots\varphi_s(\varepsilon_s)\,d\varepsilon_1\,d\varepsilon_2\ldots d\varepsilon_s.$$

Si l'on fait la somme des valeurs de cette expression pour toutes les valeurs de ε satisfaisant à la relation

$$(5) \qquad \alpha - l < \Sigma\lambda\varepsilon < \alpha + l,$$

on obtiendra la probabilité cherchée P (pour plus de dé-

tails, *voir* le raisonnement fait à la page 144); cette probabilité est alors représentée par l'intégrale

$$(6) \qquad P = \int\int \ldots \varphi_1(\varepsilon_1)d\varepsilon_1\, \varphi_2(\varepsilon_2)d\varepsilon_2 \ldots,$$

étendue aux valeurs de x satisfaisant à la relation (5). Si l'on observe que

$$\frac{1}{\pi}\int_{-\infty}^{+\infty} e^{u(\Sigma\lambda\varepsilon - \alpha)\sqrt{-1}}\,\frac{\sin ul}{u}\,du$$

est nul quand les formules (5) n'ont pas lieu, et est égal à 1 quand elles ont lieu, on pourra écrire la formule (6) ainsi :

$$P = \frac{1}{\pi}\int_{-\infty}^{+\infty} \cdots \prod\left[\varphi(\varepsilon)e^{\lambda\varepsilon u\sqrt{-1}}d\varepsilon\right]e^{-\alpha u\sqrt{-1}}\,\frac{\sin ul}{u}\,du,$$

ou bien encore

$$(7) \quad P = \frac{1}{\pi}\int_{-\infty}^{+\infty} \prod\left[\int_{-\infty}^{+\infty} \varphi(\varepsilon)e^{\lambda\varepsilon u\sqrt{-1}}d\varepsilon\right]e^{-\alpha u\sqrt{-1}}\,\frac{\sin ul}{u}\,du.$$

Posons maintenant

$$\int_{-\infty}^{\infty} \varphi_i(\varepsilon)d\varepsilon = M_0^{(i)}, \quad \int_{-\infty}^{+\infty} \varepsilon\varphi_i(\varepsilon)d\varepsilon = M_1^{(i)}, \ldots,$$

nous pourrons écrire

$$\int_{-\infty}^{+\infty} \varphi_i(\varepsilon)e^{\lambda_i\varepsilon u\sqrt{-1}}d\varepsilon = \int_{-\infty}^{+\infty} \varphi_i(\varepsilon)d\varepsilon \times \left(1 + \lambda_i\varepsilon u\sqrt{-1} - \lambda_i^2\frac{\varepsilon^2 u^2}{2} - \ldots\right)$$
$$= M_0^{(i)} + \lambda_i u M_1^{(i)}\sqrt{-1} - \frac{1}{2}\lambda_i^2 u^2 M_2^{(i)} - \ldots;$$

mais, si l'on observe que $M_0^{(i)}$ est égal à l'unité (p. 143), on pourra encore écrire

$$\int_{-\infty}^{+\infty} \varphi_i(\varepsilon) e^{\lambda_i \varepsilon u \sqrt{-1}} d\varepsilon = e^x,$$

x désignant le logarithme de

$$1 + \lambda_i u M_1^{(i)} \sqrt{-1} - \frac{1}{2} \lambda_i^2 u^2 M_2^{(i)} \ldots;$$

d'où l'on conclut

$$x = \lambda_i u M_1^{(i)} \sqrt{-1} - \lambda_i^2 u^2 \left[\frac{1}{2} M_2^{(i)} - \frac{1}{2} (M_1^{(i)})^2 \right] + \ldots;$$

et, par suite, (7) pourra s'écrire, en négligeant les termes du troisième ordre,

$$P = \frac{1}{\pi} \int_{-\infty}^{+\infty} e^{(\Sigma\lambda M_1 - \alpha)\sqrt{-1}u} e^{-\Sigma \frac{\lambda^2 u^2}{2}(M_2 - M_1^2)} \frac{\sin ul}{u} du.$$

Posons alors

(8) $$\alpha = \Sigma\lambda M_1$$

et nous aurons

$$P = \frac{1}{\pi} \int_{-\infty}^{+\infty} e^{-\Sigma \frac{\lambda^2 u^2}{2}(M_2 - M_1^2)} \frac{\sin ul}{u} du.$$

La valeur de cette intégrale est (*voir* p. 34)

(9) $$P = \frac{2}{\sqrt{\pi}} \int_0^{\gamma} e^{-\gamma^2} d\gamma,$$

formule où l'on a posé, pour abréger,

(10) $$\gamma = \frac{l}{\sqrt{2\Sigma(M_2 - M_1^2)\lambda^2}}.$$

Or M_1^2 et M_2 sont très-petits et l n'a pas besoin d'être très-grand pour que γ acquière des valeurs voisines de 2, 4 ou 3; on voit alors que P sera très-voisin de l'unité et il y aura une très-grande probabilité que ξ différera peu de $\Sigma\lambda M_1$, qui sera sa valeur la plus probable; si l'on a éliminé avec soin les erreurs systématiques, M_1 sera nul et P sera maximum, quand γ sera le plus grand possible; or γ dans ce cas est égal à $\frac{l}{\sqrt{2\Sigma M_2 \lambda^2}}$ et, pour un même écart l, γ sera d'autant plus grand, P sera d'autant plus voisin de l'unité que $\Sigma M_2 \lambda^2$ sera plus petit.

Tous les efforts du calculateur devront alors tendre à choisir pour $\lambda_1, \lambda_2, \ldots$ un système de valeurs telles que $\Sigma M_2 \lambda^2$ soit un minimum.

Avant d'aller plus loin, discutons les résultats auxquels nous venons de parvenir. Nous avons beaucoup abrégé le raisonnement qui précède, à cause de sa ressemblance en tous points avec celui de la page 144; mais nous croyons devoir insister sur ce que, si nous avons négligé les termes contenant les $M_3, M_4, \ldots$, c'est non-seulement à cause de la petitesse de ces quantités, mais, surtout encore, parce que, dans la formule (7), l'intégrale, prise par rapport à u, n'avait de valeur sensible que pour de petites valeurs de u, à cause du grand nombre de facteurs placés sous le signe Π, qui sont tous moindres que l'unité, quand u est différent de zéro, ce dont on s'assure en remarquant que déjà

$$\int_{-\infty}^{+\infty} \varphi(\varepsilon)d\varepsilon = 1.$$

Du reste, il est important d'observer que, si l'on avait posé

$$\varphi_1(\varepsilon_1) = \frac{h_1}{\sqrt{\pi}} e^{-h_1^2\varepsilon_1^2}, \quad \varphi_2(\varepsilon_2) = \frac{h_2}{\sqrt{\pi}} e^{-h_2^2\varepsilon_2^2}, \ldots,$$

c'est-à-dire si l'on avait supposé la loi de facilité des erreurs donnée par la formule de Gauss, on aurait eu au lieu de (7)

$$P = \frac{1}{\pi}\int_{-\infty}^{+\infty}\prod\left(\int_{-\infty}^{+\infty}\frac{h}{\sqrt{\pi}}e^{-h^2\varepsilon^2+\lambda\varepsilon u\sqrt{-1}}\,d\varepsilon\right)e^{-\alpha u\sqrt{-1}}\,\frac{\sin ul}{u}\,du,$$

c'est-à-dire exactement

$$P = \frac{1}{\pi}\int_{-\infty}^{+\infty}\prod e^{-\frac{\lambda^2 u^2}{4h^2}}\,e^{-\alpha u\sqrt{-1}}\,\frac{\sin ul}{u}\,du,$$

et, en prenant $\alpha = 0$,

$$P = \frac{1}{\pi}\int_{-\infty}^{+\infty} e^{-u^2\sum\frac{\lambda^2}{4h^2}}\,\frac{\sin ul}{u}\,du.$$

La valeur de cette intégrale a déjà été trouvée plusieurs fois et l'on a

$$(11)\qquad P = \frac{2}{\sqrt{\pi}}\int_0^{\gamma} e^{-\gamma^2}d\gamma,\quad \gamma = \frac{l}{\sqrt{\sum\frac{\lambda^2}{h^2}}};$$

on est ainsi amené à rendre minimum la quantité $\sum\frac{\lambda^2}{h^2}$; si l'on se rappelle que h^2 est égal à $\frac{1}{2M_2}$, on est, comme tout à l'heure, conduit à rendre minimum la quantité $\Sigma\lambda^2 M_2$ et ce résultat, qui n'était qu'approché, devient rigoureux dans l'hypothèse de Gauss sur la forme de la fonction φ.

En résumé, la question est ramenée à calculer $\lambda_1, \lambda_2, \ldots$ de telle sorte que $\sum\frac{\lambda^2}{h^2}$ soit minimum et de telle sorte aussi que les formules (3)

$$(3)\qquad \Sigma a\lambda = 1,\quad \Sigma b\lambda = 0,\quad \Sigma c\lambda = 0, \ldots$$

soient satisfaites; puis on en conclura

$$(2) \qquad x = \Sigma k\lambda,$$

avec une probabilité que l'erreur sera moindre que l, donnée par la formule (11).

Pour résoudre la question, il suffit de différentier $\sum \frac{\lambda^2}{h^2}$ et d'égaler le résultat à zéro, ce qui donne

$$\frac{\lambda_1}{h_1^2} d\lambda_1 + \frac{\lambda_2}{h_2^2} d\lambda_2 + \ldots + \frac{\lambda_s}{h_s^2} d\lambda_s = 0.$$

On a aussi, en différentiant les formules (3),

$$a_1 d\lambda_1 + a_2 d\lambda_2 + \ldots + a_s d\lambda_s = 0,$$
$$b_1 d\lambda_1 + b_2 d\lambda_2 + \ldots + b_s d\lambda_s = 0,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

Employant alors la méthode des multiplicateurs, on est conduit à poser

$$(\alpha) \qquad \left\{ \begin{array}{l} \frac{\lambda_1}{h_1^2} = \mathrm{A}a_1 + \mathrm{B}b_1 + \mathrm{C}c_1 + \ldots, \\ \frac{\lambda_2}{h_2^2} = \mathrm{A}a_2 + \mathrm{B}b_2 + \mathrm{C}c_2 + \ldots, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{array} \right.$$

L'élimination de A, B,... entre ces équations déterminera une série de relations qui, jointes aux équations (3), feront connaître les valeurs de λ. Pour faire l'élimination, multiplions la première des équations (α) par $h_1^2 a_1$, la seconde par $h_2^2 a_2$, etc.; ajoutons, en ayant égard à $\Sigma a\lambda = 1$. Cela fait, multiplions encore la première des équations (α) par $h_1^2 b_1$, la seconde par $h_2^2 b_2$, etc.; ajoutons, en ayant égard à

$\Sigma b\lambda = 0$, etc., nous aurons les formules

$$\begin{aligned} 1 &= A\Sigma h^2 a^2 + B\Sigma h^2 ab + \ldots, \\ 0 &= A\Sigma h^2 ab + B\Sigma h^2 b^2 + \ldots, \\ 0 &= A\Sigma h^2 ac + B\Sigma h^2 bc + \ldots, \\ &\ldots\ldots\ldots\ldots\ldots\ldots, \end{aligned}$$

d'où nous tirons pour A, B,... des expressions de la forme

$$(12) \qquad A = \frac{1}{\Delta}\Delta_a, \quad B = \frac{1}{\Delta}\Delta_b, \quad C = \frac{1}{\Delta}\Delta_c, \ldots,$$

où Δ représente le déterminant des coefficients de A, B, C,... et Δ_a le déterminant mineur de Δ, que l'on obtient en ôtant la première ligne et la première colonne, etc. Les formules (α) font alors connaître $\lambda_1, \lambda_2, \ldots$; mais, si l'on forme la quantité $\Sigma\lambda k$, qui est égale à x, en vertu de l'équation (2), on trouve

$$x = A\Sigma h^2 ak + B\Sigma h^2 bk + C\Sigma h^2 ck + \ldots$$

et, en remplaçant A, B, C,... par leurs valeurs (12),

$$x = \frac{1}{\Delta}(\Delta_a \Sigma h^2 ak + \Delta_b \Sigma h^2 bk + \Delta_c \Sigma h^2 ck + \ldots).$$

Or il est important d'observer que la valeur de x ainsi obtenue, que les valeurs de y, z,... obtenues d'une façon analogue, sont les solutions des équations

$$(13) \quad \begin{cases} x\Sigma h^2 a^2 + y\Sigma h^2 ab + z\Sigma h^2 ac + \ldots = \Sigma h^2 ak, \\ x\Sigma h^2 ab + y\Sigma h^2 b^2 + z\Sigma h^2 cb + \ldots = \Sigma h^2 bk, \\ \ldots\ldots\ldots\ldots\ldots\ldots \end{cases}$$

Gauss parvient à ces mêmes formules d'une autre manière, en observant que la probabilité du concours des er-

reurs $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_s$ est

$$h_1 h_2 \ldots h_s \pi^{\frac{s}{2}} e^{-\Sigma h^2 \varepsilon^2} d\varepsilon_1\, d\varepsilon_2 \ldots d\varepsilon_s$$

et que le système d'erreurs le plus probable correspond au maximum de cette expression, c'est-à-dire au minimum de $\Sigma h^2 \varepsilon^2$. Il est facile de prouver que les valeurs de x, y, $z, \ldots$, qui rendent $\Sigma \varepsilon^2 h^2$ minimum, satisfont aux formules (13). En effet, on a

$$(14)\qquad \left\{\begin{array}{l} a_1 x + b_1 y + c_1 z + \ldots - k_1 = \varepsilon_1, \\ a_2 x + b_2 y + c_2 z + \ldots - k_2 = \varepsilon_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \end{array}\right.$$

et, par suite,

$$\Sigma h^2 \varepsilon^2 = h_1^2 (a_1 x + b_1 y + \ldots - k_1)^2 + h_2^2 (a_2 x + b_2 y + \ldots - k_2)^2 + \ldots.$$

Pour trouver le minimum de cette expression, il suffit d'égaler à zéro ses dérivées partielles, par rapport à x, y, $z, \ldots$, ce qui donne

$$\begin{array}{l} a_1 h_1^2 (a_1 x + b_1 y + \ldots - k_1) + a_2 h_2^2 (a_2 x + b_2 y + \ldots - k_2) + \ldots = 0, \\ b_1 h_1^2 (a_1 x + b_1 y + \ldots - k_1) + b_2 h_2^2 (a_2 x + b_2 y + \ldots - k_2) + \ldots = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{array}$$

Si l'on ordonne ces équations, par rapport à x, y, $z, \ldots$, on retombe sur les formules (13). Lorsque les observations comportent la même précision, ce que l'on suppose ordinairement, la méthode revient, comme on voit, à rendre $\Sigma \varepsilon^2$ minimum; de là le nom de *Methode des moindres carrés* donné au procédé que nous venons d'indiquer.

En résumé, la méthode des moindres carrés suppose le nombre des équations à résoudre très-grand; elle n'est rigoureuse qu'à cette condition. Il est important, en effet, de remarquer que la forme assignée par Gauss à la fonction φ

repose sur cette hypothèse, que la moyenne arithmétique d'un système de valeurs approchées d'une quantité est la valeur la plus probable de cette quantité (p. 154); or il serait tout aussi naturel d'admettre que la moyenne géométrique, que la moyenne harmonique, etc., représentent la valeur la plus probable de la quantité cherchée, et à chacune de ces hypothèses correspondrait une forme différente de φ.

Pour appliquer la méthode des moindres carrés, et surtout pour pouvoir calculer la probabilité P que l'erreur ξ de x ne dépasse pas l, il faudrait connaître les quantités M_2 ou h; or les rapports de ces quantités à l'une quelconque d'entre elles sont connus. En effet, s'il n'y a aucune raison d'attacher plus d'importance à une observation qu'à une autre, on est bien obligé de supposer les fonctions $\varphi_1, \varphi_2, \ldots$ identiques, et, par suite,

$$M_2^{(1)} = M_2^{(2)} \ldots \quad \text{ou} \quad h_1^2 = h_2^2 \ldots .$$

Si, au contraire, il y a quelque raison plausible pour attacher des poids inégaux aux diverses observations $k_1, k_2, \ldots$, on connaîtra généralement la mesure de la précision relative de ces observations. Par exemple, $k_1, k_2, \ldots$ pourront être des angles mesurés respectivement par $p_1, p_2, \ldots$ répétitions, et les quantités $M_2^{(1)}, M_2^{(2)}, \ldots$ pourront être grossièrement calculées en s'appuyant sur les résultats donnés (p. 152 et suiv.); dans d'autres circonstances, il sera plus difficile de déterminer les rapports $M_2^{(1)} : M_2^{(2)} \ldots$, et, en l'absence de toute règle, c'est par le sentiment qu'on devra se laisser guider pour fixer les valeurs de ces rapports.

Quoi qu'il en soit, supposons les rapports $M_2^{(1)} : M_2^{(2)} \ldots$ ou $h_1^2 : h_2^2 \ldots$ connus et désignons-les par $\theta_1 : \theta_2 \ldots$, les formules (13) deviendront

$$x\Sigma\theta a^2 + y\Sigma\theta ab + \ldots = \Sigma\theta k,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

et permettront de calculer la valeur la plus probable de x ou $\Sigma\lambda M_1$. Supposons que l'on connaisse les valeurs les plus probables de y, de $z, \ldots$; désignons par $x', y', z', \ldots$ ces valeurs les plus probables, la différence entre

$$a_i x' + b_i y' + c_i z' \ldots \quad \text{et} \quad k_i$$

sera à peu près égale à ε_i. Or il est facile de prouver, par une analyse toute semblable à celle qui précède, que $\Sigma\lambda^2 M_2$ est très-probablement fort peu différent de $\Sigma\lambda^2\varepsilon^2$; les quantités λ sont connues, les ε peuvent être calculés approximativement, comme nous venons de le dire, et, par suite, la probabilité P donnée par la formule (11) pourra être calculée avec une exactitude très-suffisante.

Nous terminerons ce paragraphe par quelques remarques importantes :

1° Si l'on applique la méthode des moindres carrés à des équations à une seule inconnue, dans le cas où la précision de chaque observation reste la même, il est facile de voir qu'elle conduit à prendre pour valeur de l'inconnue la moyenne des observations. En effet, les équations

$$x = k_1, \quad x = k_2, \ldots, \quad x = k_s$$

donnent l'équation unique

$$\Sigma x = \Sigma k \quad \text{ou} \quad x = \frac{\Sigma k}{s}.$$

2° M. Yvon Villarceau a déjà fait remarquer que des méthodes en apparence très-différentes de celle des moindres carrés pouvaient conduire à des résultats peu différents de ceux que l'on obtiendrait en suivant cette voie. Si les nouvelles méthodes paraissent très-expéditives, on peut les appliquer, et la théorie exposée ci-dessus permettra de calculer la probabilité des résultats que l'on trouve. En

effet, les multiplicateurs λ ne sont soumis à aucune hypothèse particulière dans le calcul de P, et toute méthode d'élimination revient au fond, comme on le sait, à la méthode des multiplicateurs.

Généralisation de la méthode des moindres carrés.

Lorsque les équations que l'on a à résoudre n'ont pas la forme linéaire, on peut encore leur appliquer la méthode des moindres carrés et voici comment. Soient

$$(1)\quad f_1(x,y,z,\ldots)=0,\quad f_2(x,y,z,\ldots)=0,\ldots,\quad f_s(x,y,z,\ldots)=0$$

les équations à résoudre, on commencera par calculer, à l'aide de n de ces équations, les n inconnues $x, y, z, \ldots$; les valeurs trouvées $\alpha, \beta, \gamma, \ldots$ seront des valeurs approchées de $x, y, z, \ldots$; en posant alors

$$x=\alpha+\delta x,\quad y=\beta+\delta y,\quad z=\gamma+\delta z,\ldots,$$

les formules (1) deviendront

$$f_1+\frac{\partial f_1}{\partial \alpha}\delta x+\frac{\partial f_1}{\partial \beta}\delta y+\ldots=0,$$
$$f_2+\frac{\partial f_2}{\partial \alpha}\delta x+\frac{\partial f_2}{\partial \beta}\delta y+\ldots=0,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,$$

et l'on pourra négliger les termes du deuxième ordre en δx, $\delta y, \ldots$. Ces équations sont alors du premier degré et on peut leur appliquer la méthode des moindres carrés; cette méthode conduit alors à évaluer de très-petites valeurs, qui devront être considérées comme des termes de correction.

Remarques de M. Bienaymé.

Il y a quelques années, M. Bienaymé attira l'attention des géomètres sur ce point important, que la méthode des moindres carrés, telle qu'elle se trouvait exposée dans les Ouvrages classiques, était loin de fournir les valeurs les plus probables des inconnues avec toute la précision qu'on lui avait attribuée tout d'abord. Et, en effet, on avait autrefois l'habitude de calculer séparément la probabilité de la valeur de chaque inconnue, comme si elle eût été indépendante des autres; on était loin d'obtenir ainsi la probabilité du système adopté pour l'ensemble des inconnues.

Si l'on désigne par p la probabilité que l'inconnue x sera comprise entre $\alpha - l$ et $\alpha + l$, par q la probabilité que l'inconnue y sera comprise entre $\beta - m$ et $\beta + m, \ldots$, on voit que la probabilité du concours de ces circonstances sera à peu près $pq\ldots.$ Je dis à peu près, parce que ces circonstances ne sont pas tout à fait indépendantes les unes des autres, et il convient de donner à cet égard une analyse rigoureuse; on voit cependant, par le raisonnement qui précède, que la probabilité du système des valeurs admis pour $x, y, \ldots$ sera moindre que la probabilité de la valeur attribuée à chaque inconnue. Toutefois il est bon d'observer que les méthodes exposées plus haut sont suffisantes, tant que l'on ne désire pas calculer toutes les inconnues, mais seulement une seule d'entre elles. Enfin il faut rappeler que Laplace, en justifiant pour la première fois la méthode des moindres carrés, sans faire d'hypothèse sur la forme de la fonction φ, ne s'est pas rendu coupable des fautes que M. Bienaymé reproche aux géomètres plus modernes. Nous ne pouvons pas ici reproduire complétement l'analyse donnée par M. Bienaymé, dans le *Journal de*

M. Liouville, t. XVII; mais, profitant de ces remarques, nous donnerons une solution approchée de la question en suivant une autre méthode.

Supposons que les équations à résoudre soient, comme plus haut,

$$(1)\quad \begin{cases} a_1 x + b_1 y + c_1 z \ldots = k_1, \\ a_2 x + b_2 y + c_2 z \ldots = k_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_s x + b_s y + c_s z \ldots = k_s. \end{cases}$$

Nous supposerons que le nombre des inconnues $x, y, z, \ldots$ soit r. Pour résoudre le système (1), on fait, comme on le sait, usage de la méthode des multiplicateurs, et l'on pose

$$(2)\quad \begin{cases} \Sigma a\lambda = 1, & \Sigma b\lambda = 0, & \Sigma c\lambda = 0, \ldots, \\ \Sigma a\mu = 0, & \Sigma b\mu = 1, & \Sigma c\mu = 0, \ldots, \\ \Sigma a\nu = 0, & \Sigma b\nu = 0, & \Sigma c\nu = 1, \ldots, \\ \ldots\ldots\ldots, & \ldots\ldots\ldots, & \ldots\ldots\ldots\ldots, \end{cases}$$

le signe Σ s'étendant aux valeurs des indices $1, 2, 3, \ldots, s$ des $a, b, c, \ldots$ et des $\lambda, \mu, \nu, \ldots$; on choisit ensuite

$$(3)\qquad x = \Sigma\lambda k, \quad y = \Sigma\mu k, \quad z = \Sigma\nu k, \ldots.$$

Les formules (2) ne suffisant pas à la détermination des $\lambda, \mu, \nu, \ldots$, puisqu'elles ne sont qu'au nombre de r^2, on achève de déterminer ces quantités en leur imposant la condition de donner au système $x, y, z, \ldots$ un maximum de probabilité.

Pour y parvenir, appelons $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_s$ les erreurs des observations $k_1, k_2, \ldots, k_s$ et $\xi, \eta, \zeta, \ldots$ les erreurs qui en résultent pour $x, y, z, \ldots$, les valeurs exactes $x + \xi$, $y + \eta, \ldots$ devant satisfaire aux formules (1) quand on y

remplace k par $k+\varepsilon$, on a

$$x+\xi=\Sigma\lambda k+\Sigma\lambda\varepsilon,\quad y+\eta=\Sigma\mu k+\Sigma\mu\varepsilon,\ldots;$$

de ces formules et des formules (3) on tire

$$(4)\qquad \xi=\Sigma\lambda\varepsilon,\quad \eta=\Sigma\mu\varepsilon,\quad \zeta=\Sigma\nu\varepsilon,\ldots.$$

Soit maintenant $\varphi_i(\varepsilon_i)$ la facilité de l'erreur ε_i commise sur l'observation k_i et, pour plus de généralité, supposons les facilités des erreurs différentes dans les diverses observations : l'observation k_1 sera, par exemple, un angle mesuré par p_1 répétitions, k_2 sera un angle mesuré par p_2 répétitions, et ainsi de suite;

$$(5)\qquad \varphi_1(\varepsilon_1)d\varepsilon_1\,\varphi_2(\varepsilon_2)d\varepsilon_2\ldots\varphi_s(\varepsilon_s)d\varepsilon_s$$

sera la probabilité du concours des erreurs $\varepsilon_1, \varepsilon_2, \ldots$ sur $k_1, k_2, \ldots$, ou d'erreurs différant de celles-là de quantités inférieures à $d\varepsilon_1, d\varepsilon_2, \ldots, d\varepsilon_s$. En intégrant ce produit de telle sorte que l'on ait

$$(6)\qquad \sqrt{\frac{(\xi-\alpha)^2}{l^2}+\frac{(\eta-\beta)^2}{m^2}+\frac{(\zeta-\gamma)^2}{n^2}\cdots}<1,$$

on aura la probabilité que les erreurs des inconnues satisferont à cette relation. Pour effectuer le calcul, nous multiplierons la quantité à intégrer par

$$\frac{1}{\pi}\int_{-\infty}^{+\infty}\cos G u\,\frac{\sin u}{u}\,du,$$

qui est égal à zéro ou à l'unité, selon que G est plus petit ou plus grand que l'unité, et nous poserons

$$G=\sqrt{\left(\frac{\xi-\alpha}{l}\right)^2+\left(\frac{\eta-\beta}{m}\right)^2+\ldots}$$

La présence du facteur que nous introduisons ainsi nous permettra de prendre, pour limites des ε, $-\infty$ et $+\infty$. En désignant alors par P la probabilité cherchée, nous aurons

$$P = \frac{1}{\pi}\int_{-\infty}^{+\infty} \cdots \varphi_1(\varepsilon_1)\ldots\varphi(\varepsilon_s)d\varepsilon_1\ldots d\varepsilon_s \cos G u \sin u \frac{du}{u}.$$

Si nous développons $\cos G u$ suivant les puissances ascendantes de u, nous trouvons

$$P = \frac{1}{\pi}\int_{-\infty}^{+\infty} \cdots \varphi_1(\varepsilon_1)\ldots\varphi_s(\varepsilon_s)d\varepsilon_1\ldots d\varepsilon_s$$
$$\times\left(1 - \frac{G^2u^2}{1.2} + \frac{G^4u^4}{4!} - \ldots\right)\sin u \frac{du}{u}.$$

Or G^ω, en général, est une fonction entière des ε; en effet, G^ω est fonction entière des $\xi, \eta, \zeta, \ldots$ respectivement égaux à $\Sigma\lambda\varepsilon$, $\Sigma\mu\varepsilon$, $\Sigma\nu\varepsilon$, ... Si alors on pose

$$\int_{-\infty}^{+\infty} \varphi_1(\varepsilon_1)\varepsilon_1^\omega d\varepsilon = M_1^\omega, \quad \int_{-\infty}^{+\infty} \varphi_2(\varepsilon_2)\varepsilon_2^\omega d\varepsilon_2 = M_2^\omega, \ldots,$$

l'intégrale

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \cdots G^i\varphi(\varepsilon_1)\ldots\varphi(\varepsilon_s)d\varepsilon_1\ldots d\varepsilon_s$$

pourra être représentée par Γ_i, en désignant ainsi ce que devient G^i, quand on y remplace ε_1^ω par M_1^ω, ε_2^ω par $M_2^\omega, \ldots$, et l'on trouve alors

$$P = \frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{\sin u}{u}\left(1 - \frac{\Gamma_2 u^2}{2!} + \frac{\Gamma_4 u^4}{4!} - \ldots\right)du,$$

ou bien, en négligeant les termes du quatrième ordre et en

remplaçant les termes conservés par une exponentielle,

(7) $$P = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\sin u}{u} e^{-\Gamma_2 \frac{u^2}{2}} du,$$

c'est-à-dire (p. 34)

(8) $$P = \frac{2}{\sqrt{\pi}} \int_0^{\frac{1}{\sqrt{2\Gamma_2}}} e^{-t^2} dt.$$

Il reste à évaluer la constante Γ_2.

On a

$$G^2 = \left(\frac{\Sigma\lambda\varepsilon - \alpha}{l}\right)^2 + \left(\frac{\Sigma\mu\varepsilon - \beta}{m}\right)^2 + \ldots,$$

ou

$$G^2 = \mathbf{S}\left(\frac{\Sigma\lambda\varepsilon - \alpha}{l}\right)^2,$$

c'est-à-dire

$$G^2 = \mathbf{S}\, \frac{\Sigma\lambda^2\varepsilon^2 + 2\Sigma\lambda_i\lambda_j\varepsilon_i\varepsilon_j - 2\alpha\Sigma\lambda\varepsilon + \alpha^2}{l^2},$$

et, par suite,

$$\Gamma_2 = \mathbf{S}\, \frac{1}{l^2}\left[\Sigma\lambda^2 M^{(2)} + 2\Sigma\lambda_i\lambda_j M_i^{(1)} M_j^{(1)} - 2\alpha\Sigma\lambda M^{(1)} + \alpha^2\right].$$

Cette formule se simplifie en posant

(9) $$\alpha = \Sigma\lambda M^{(1)}, \quad \beta = \Sigma\mu M^{(1)}, \ldots;$$

en effet, elle devient ainsi

$$\Gamma_2 = \mathbf{S}\, \frac{1}{l^2}\left[\Sigma\lambda^2 M^{(2)} + 2\Sigma\lambda_i\lambda_j M_i^{(1)} M_j^{(1)} - 2(\Sigma\lambda M^{(1)})^2 + (\Sigma\lambda M^{(1)})^2\right],$$

c'est-à-dire

$$r_2 = \mathrm{S} \sum \frac{\lambda^2}{l^2} [M^{(2)} - (M^{(1)})^2].$$

La formule (8) devient alors

$$(10) \qquad P = \Theta \left[\frac{1}{\sqrt{2 \sum [M^{(2)} - (M^{(1)})^2] \left(\frac{\lambda^2}{l^2} + \frac{\mu^2}{m^2} + \frac{\nu^2}{n^2} + \ldots \right)}} \right],$$

Θ désignant, pour abréger, un signe de fonction défini par la formule

$$\Theta(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2} dt.$$

P sera très-voisin de l'unité dès que la quantité placée sous le signe Θ atteindra une valeur telle que 3, comme l'on peut s'en assurer en consultant une table de la fonction Θ; on est ainsi conduit, pour rendre P très-voisin de l'unité, à rendre les sommes $\Sigma M^{(2)} \lambda^2$, $\Sigma M^{(2)} \mu^2, \ldots$ aussi petites que possible, en supposant les erreurs constantes éliminées. La méthode des moindres carrés se trouve donc maintenant établie d'une manière générale. Quant au calcul de P et des M_2, il se fera comme nous l'avons indiqué plus haut. (*Voir* la Note III placée à la fin du volume.)

Application de la théorie précédente à l'appréciation de la justesse d'une arme à feu.

Je vais maintenant montrer comment on devrait procéder pour apprécier la justesse des armes à feu; cet exemple me paraît fort bien choisi pour faire ressortir l'utilité des théories qui précèdent.

Poisson s'était déjà occupé, en 1837, de cette question,

dans un Mémoire qui a été inséré au *Mémorial de l'Officier d'Artillerie ;* la méthode que Poisson a proposée ne diffère pas essentiellement de celle que nous allons développer, en ce sens qu'elle conduit aux mêmes conclusions, mais son analyse diffère notablement de la nôtre. Le général Didion s'est aussi occupé de cette question et le Traité qu'il a écrit peut être consulté avec fruit.

Nous appellerons points d'*impact* les points où les projectiles viennent atteindre une cible, sur laquelle nous supposerons que l'on ait tiré un certain nombre de coups avec une même arme.

Soit O le centre de la cible ou le point sur lequel on vise; par ce point, traçons deux axes sur la cible, l'un horizontal, l'autre vertical, et soient x et y les coordonnées d'un point d'impact. Désignons par $\varphi(x)dx$ la probabilité que l'abscisse du point d'impact sera comprise entre x et $x+dx$ et par $\psi(y)dy$ la probabilité que l'ordonnée du même point sera comprise entre y et $y+dy$. Si l'arme que l'on veut éprouver ne possède en elle-même aucune cause qui tende à faire dévier le projectile, on devra avoir

$$\varphi(x)=\varphi(-x) \quad \text{et} \quad \psi(y)=\psi(-y);$$

en tout cas, on aura

$$(1) \qquad \int_{-\infty}^{+\infty}\varphi(x)dx=1, \quad \int_{-\infty}^{+\infty}\psi(y)dy=1.$$

En effet, les deux intégrales précédentes représentent les probabilités que le point d'impact se trouvera en un point quelconque de la cible, supposée indéfinie : du reste $\varphi(x)$ et $\psi(y)$ décroissent rapidement, quand x et y vont en croissant en valeur absolue, x et y pouvant en quelque sorte être considérés comme des erreurs dans la position du point

12.

d'impact, qui devrait être l'origine des coordonnées, si l'arme était parfaite et le pointeur d'une adresse irréprochable.

Ceci posé, cherchons la probabilité que, sur un grand nombre s de coups, la somme des abscisses des points d'impact sera comprise entre $\alpha - l$ et $\alpha + l$. En nous reportant à la théorie développée au § II, on voit que cette probabilité est donnée par la formule

$$(2) \qquad P = \frac{2}{\sqrt{\pi}} \int_0^{\gamma} e^{-\gamma^2} d\gamma,$$

dans laquelle

$$\gamma = \frac{l}{\sqrt{2s(M_2 - M_1^2)}},$$

et où l'on a choisi $\alpha = M_1 s$, M_1 et M_2 désignant toujours, comme plus haut, les intégrales

$$M_1 = \int_{-\infty}^{+\infty} x\varphi(x)dx, \quad M_2 = \int_{-\infty}^{+\infty} x^2\varphi(x)dx.$$

La probabilité que la moyenne des abscisses sera comprise entre $M_1 - \varepsilon$ et $M_1 + \varepsilon$ sera donnée par la formule (2), où l'on changera l en εs, ou par

$$(3) \qquad \begin{cases} P = \dfrac{2}{\sqrt{\pi}} \displaystyle\int_0^{\gamma} e^{-\gamma^2} d\gamma, \\ \gamma = \dfrac{\varepsilon\sqrt{s}}{\sqrt{2(M_2 - M_1^2)}}. \end{cases}$$

Ceci posé, lorsque s sera assez grand, γ deviendra assez grand pour que P soit très-voisin de l'unité, mais alors le point d'impact le plus probable est celui qui a pour abscisse M_1. Supposons que l'on ait tiré un certain nombre de

coups dans la cible avec une arme donnée : en faisant la moyenne des abscisses des points d'impact observés, si l'on trouve un résultat qui diffère notablement de zéro, on en conclura que M_1 n'est pas nul et que, par suite, $\varphi(x)$ n'est pas égal à $\varphi(-x)$: il existera donc une cause constante qui tend à faire dévier le projectile du plan de tir. Pour savoir quel est l'écart que l'on peut raisonnablement attribuer au hasard, on aura recours aux formules (3) ; on y remplacera M_1 par la moyenne des abscisses des points d'impact, et M_2 par la moyenne de leurs carrés, en se donnant alors P très-voisin de l'unité et égal à 0,9999, par exemple ; il faudra prendre $\gamma = 3,1$, et l'on en conclura ε, ce qui fera connaître en quelque sorte des limites que l'abscisse du point d'impact ne dépassera probablement pas. Nous rappellerons, du reste, que, sur un très-grand nombre de coups, la moyenne des carrés des abscisses des points d'impact est sensiblement M_2 (*voir* p. 152).

Ce que nous venons de dire des écarts latéraux peut se répéter sur les écarts en hauteur, car on peut répéter sur la fonction ψ ce que nous avons dit de la fonction φ, mais l'épreuve dont nous venons de parler n'est pas suffisante pour faire apprécier la justesse d'une arme à feu ; en effet, cette épreuve, indispensable à faire, a pour but seulement de constater la présence des *erreurs systématiques dans l'arme* ou, si l'on veut, des causes constantes qui tendent à écarter les projectiles du centre de la cible. Il pourrait fort bien se faire que ces causes constantes n'existassent point et que l'arme fût mauvaise, à cause des grands écarts qu'elle pourrait donner autour du but à atteindre.

Pour bien juger une arme, il faudrait donc avoir encore une idée des écarts qu'elle peut donner autour du but ; or, en supposant M_1 très-petit, ce qui est une première condition de justesse, la formule (3) montre que plus M_2 sera

petit et plus γ sera grand, et, par suite aussi, plus P sera voisin de l'unité; mais M_2, dans un grand nombre de coups, est sensiblement égal à la moyenne des carrés des abscisses des points d'impact; donc, enfin, la meilleure arme, celle qui donnera les moindres écarts latéraux, sera celle pour laquelle la moyenne des carrés des abscisses des points d'impact sera un minimum. Il y aurait une observation analogue à faire sur les écarts en hauteur.

En résumé : 1° pour que l'arme ne présente pas de cause tendant à faire dévier le projectile, il faut que la moyenne des abscisses et des ordonnées des points d'impact soit très-petite, ou, si l'on veut, que le centre de gravité de ces points d'impact soit voisin du centre de la cible : la formule (3) permet d'apprécier la limite supérieure admissible pour cette distance; 2° il faut que les moyennes des carrés des abscisses et des ordonnées des points d'impact soient elles-mêmes très-petites.

Sans qu'il soit nécessaire de pousser plus loin cette discussion, on voit l'analogie qu'elle présente avec la théorie des erreurs. On voit aussi que l'on peut, dans les applications, admettre, pour la fonction φ et pour la fonction ψ, la forme suivante, où h se détermine facilement en fonction de M_2 que l'on peut observer,

$$\frac{h}{\sqrt{\pi}} e^{-h^2 x^2};$$

mais il faut pour cela que s soit très-grand; h donne alors la mesure de la précision de l'arme. Poisson, dans le Mémoire que nous avons cité, examine encore la manière dont les coups doivent se distribuer angulairement autour du centre de la cible. Pour y parvenir, il fait usage des coordonnées polaires : la discussion de ce cas est analogue en tout point à celle que nous venons de faire.

Application des théories précédentes à la recherche des lois des phénomènes.

La recherche des lois d'un phénomène se réduit, comme l'on sait, en dernière analyse, à la recherche d'une fonction, dont on connaît un grand nombre de valeurs, avec les valeurs correspondantes de la variable; mais il se présente deux cas à examiner, dont nous donnerons deux exemples. Dans l'un des cas, et c'est le plus ordinaire, on peut admettre qu'une seule observation, pour une même valeur de la variable, fournit une valeur très-approchée de la fonction : tel serait le cas où l'on aurait à étudier la manière dont varie l'allongement d'une barre avec la température. D'autres fois, au contraire, il faut un grand nombre d'observations pour fixer la valeur de la fonction correspondant à une valeur donnée de la variable. La construction des Tables de mortalité, qui indiquent combien, sur n individus, nés le même jour, il en reste en vie à l'âge x, est un exemple du second genre de questions.

Je suppose donc qu'il s'agisse de déterminer une fonction de la température t, telle que, pour $t = t_0, t_1, t_2, \ldots$, elle prenne les valeurs $\delta_0, \delta_1, \delta_2, \ldots$, qui représentent les longueurs d'une barre aux températures $t_0, t_1, t_2, \ldots$.

Pour résoudre cette question, on commencera par fixer *a priori* le degré d'exactitude sur lequel on peut compter dans chaque observation; on saura, par exemple, que la longueur δ est mesurée à $\frac{1}{n}$ près et que la température est mesurée à $\frac{1}{m}$ près, *grosso modo*. On pourra estimer qu'une variation de $\frac{1}{m}$ dans t ne produit pas une variation $\frac{1}{\mu}$ dans δ

et, par suite, pour une valeur exacte de t, on évaluera l'erreur admissible pour δ comme inférieure à $\frac{1}{n} + \frac{1}{\mu} = E$.

Les observations, dans le cas dont il s'agit, sont censées faites avec le même soin, en sorte que nous supposerons la précision constante. Nous chercherons alors à représenter la longueur δ de la barre à la température t par une expression de la forme

$$\delta = a + bt.$$

Pour cela, nous substituerons à la place de δ et de t les valeurs observées, et nous aurons des équations telles que

$$(1) \qquad \delta_0 = a + bt_0, \quad \delta_1 = a + bt_1, \ldots, \quad \delta_s = a + bt_s,$$

qu'il faudra bien se garder de simplifier dans le cas où l'on aurait des dénominateurs que l'on voudrait chasser. En considérant a et b comme des inconnues, on appliquera aux formules (1) la méthode des moindres carrés et l'on calculera la probabilité des valeurs trouvées pour a et b. Si l'on trouve cette probabilité assez voisine de l'unité, on adoptera définitivement la forme

$$\delta = a + bt;$$

si, au contraire, on trouve une faible probabilité pour a et b, on essayera la forme

$$\delta = a + bt + ct^2.$$

On déterminera a, b, c par la méthode des moindres carrés, au moyen des formules

$$\delta_0 = a + bt_0 + ct_0^2, \quad \delta_1 = a + bt_1 + ct_1^2, \ldots,$$

et l'on calculera encore la probabilité des valeurs trouvées

pour a, b, c, et ainsi de suite. Il pourra se faire qu'aucune des formes paraboliques ne donne des valeurs satisfaisantes pour les coefficients; on pourra alors essayer d'autres formes que l'on discutera de la même façon.

J'ai dit que l'on devait calculer la probabilité des valeurs de a, b, c,...; ceci exige un mot d'explication. On a vu (p. 173) qu'il y avait une probabilité

$$P = \frac{2}{\sqrt{\pi}} \int_0^{\gamma} e^{-\gamma^2} d\gamma, \quad \gamma = \frac{1}{\sqrt{2 \Sigma M_2 \left(\frac{\lambda^2}{l^2} + \frac{\mu^2}{m^2} + \ldots \right)}}$$

que les inconnues étaient entachées d'erreurs ξ, η,..., inférieures, en valeur absolue, à l, m,..., et du reste satisfaisant à l'inégalité

$$\sqrt{\frac{\xi^2}{l^2} + \frac{\eta^2}{m^2} + \ldots} < 1.$$

Or M_2, dans le cas actuel, est donné par la condition que l'erreur commise sur δ_0, δ_1,... est très-probablement moindre que E. En effet, si s est très-grand, on peut admettre que la facilité de l'erreur est donnée par la formule

$$\frac{h}{\sqrt{\pi}} e^{-h^2 u^2}$$

et l'on a

$$M_2 = \frac{1}{2h^2};$$

on pourra alors poser, par exemple,

$$\int_{-E}^{+E} \frac{h}{\sqrt{\pi}} e^{-h^2 E^2} dE = 0,999,$$

d'où l'on déduira h et par suite M_2. Connaissant M_2, on calculera P en essayant pour l, m,... diverses valeurs, et

l'on rejettera celles qui donneraient, pour les erreurs de δ_0, $\delta_1, \ldots$, calculées *a posteriori*, des valeurs supérieures à E avec une valeur de P trop petite. (Ce calcul *a posteriori* se fera en substituant à a, $b, \ldots$ dans la formule (1) leurs limites les plus défavorables.) Si l'on ne peut pas trouver de valeurs convenables pour les $l, m, n, \ldots$, il faudra rejeter la formule d'interpolation adoptée.

Dans l'exemple que nous avons choisi, on pouvait attribuer la même précision à toutes les observations; mais il n'en est pas toujours ainsi. Lorsque l'on effectue, par exemple, des mesures géodésiques, certains angles peuvent être résumés avec plus de précision que d'autres, selon qu'ils ont été obtenus par un nombre de répétitions plus grand ou avec des instruments meilleurs, selon qu'ils ont été mesurés directement ou calculés à l'aide de réseaux triangulaires, etc.; mais on pourra, dans la plupart des cas, estimer la précision et déterminer, avec une approximation suffisante, la facilité des erreurs. Nous renverrons, pour ces questions de détail, aux Traités spéciaux : les développements que comporte cette question nous entraîneraient à des digressions sortant du cadre de cet Ouvrage, qui a un caractère essentiellement théorique (*).

(*) Les erreurs d'observations suivent parfois des lois bizarres; ainsi, dans l'exemple que nous avons choisi, relatif à la dilatation d'une barre, la facilité de l'erreur ε est non-seulement une fonction de ε, mais encore une fonction de la température, car la difficulté des mesures croît avec la température. A défaut de tout autre moyen de pouvoir tenir compte du poids des observations, l'expérimentateur pourra attribuer un coefficient à chaque équation de condition : ce coefficient sera analogue à ceux que l'on a imaginés pour apprécier le mérite d'un candidat dans un examen, ce coefficient servira de facteur à l'équation à laquelle on voudra attribuer une importance en rapport avec lui. Sans doute cette façon de procéder est empirique, mais elle fournira toujours de meilleurs résultats que si l'on faisait abstraction de l'importance que certaines observations peuvent avoir.

Il faut avouer que la méthode d'interpolation que nous venons d'indiquer ne saurait être défendue par un esprit rigoureux. Nous ne la donnons (d'ailleurs sous toutes réserves) que parce qu'il peut souvent ne pas en exister de meilleure. Quoi qu'il en soit, la méthode des moindres carrés devra toujours être appliquée sans hésitation, quand on connaîtra d'avance, comme cela a lieu le plus souvent en Astronomie, la forme de la fonction interpolatrice et que ses coefficients seuls seront inconnus.

Des Tables de mortalité.

Une Table de mortalité est une Table qui fait connaître le nombre d'individus survivants à un âge déterminé, sur un nombre donné d'individus nés le même jour.

Une Table de mortalité devrait donc indiquer la probabilité qu'a l'homme d'atteindre un âge donné; en effet, si l'on sait que sur y_0 individus, nés le même jour, il en reste y à l'âge x, le rapport du nombre des cas où l'homme vit x années au nombre des cas où il vit x années et où il ne vit pas x années est $\frac{y}{y_0}$; mais ce n'est là qu'une approximation d'autant plus grande que y_0 a été choisi plus grand. Le théorème de Bayes nous apprend en effet que, pour trouver la probabilité de l'événement, qui est l'existence d'un individu à l'âge x, il faut faire un nombre y_0 d'épreuves consistant à choisir y_0 individus d'âge o et de compter combien parmi eux arrivent à l'âge x : le rapport $\frac{y}{y_0}$ tend vers la probabilité cherchée quand y_0 croît indéfiniment.

Il y a plus, le théorème de Bayes, convenablement appliqué, nous permet de déterminer l'erreur commise sur la probabilité cherchée, en prenant pour cette probabilité le

rapport $\frac{y}{y_0}$, y_0 étant un nombre fixe, mais suffisamment grand.

Toutefois, en examinant la question de plus près, on se heurte forcément contre une difficulté qui tient à la manière même dont on compte les épreuves, c'est-à-dire à la manière dont on observe les faits. Pour construire une Table de mortalité, on est obligé de s'adresser à une classe d'individus dont on peut suivre plus ou moins exactement les naissances et les décès, en sorte que les probabilités auxquelles on arrive diffèrent nécessairement suivant les classes sur lesquelles ont porté les observations ; c'est ce qui explique les divergences que l'on observe entre des Tables construites d'ailleurs avec les soins les plus minutieux. On trouve, par exemple, que la mortalité est beaucoup plus grande à Berlin et dans les capitales du monde civilisé que dans le reste de l'Europe, plus grande chez les hommes que chez les femmes, plus grande à Vienne qu'à Genève, etc.

Il serait extrêmement difficile de construire une Table pour tout le genre humain : les documents statistiques manquent d'une façon absolue. Doit-on le regretter? Si l'on se place au point de vue des applications, on peut répondre, sans hésiter, non; et, en effet, l'étude de la mortalité a surtout pour objet la recherche des causes qui peuvent tendre à diminuer la vie humaine dans diverses localités ou parmi certaines classes d'individus, et la théorie des opérations financières. Dans les deux premiers cas, il va sans dire que les lois de mortalité que l'on recherche ne doivent s'étendre qu'à une fraction déterminée du genre humain; dans le dernier cas, nous verrons qu'il est impossible de s'en rapporter à une loi générale, et que deux Tables distinctes doivent être employées suivant la nature des opérations que l'on doit faire.

Je suppose que l'on ait construit, avec une exactitude suffisante, une Table de mortalité pour une classe déterminée d'individus, le nombre des survivants y à l'âge x sera une fonction de x, que nous pourrons appeler $f(x)$, et l'équation

$$y = f(x)$$

représentera une courbe qui porte le nom de *courbe de mortalité*. La plupart des courbes de mortalité sont sans inflexions et ressemblent plus ou moins au quart de la développée de l'ellipse. La courbe de mortalité de Genève, cependant, présente un renflement vers son milieu, et, par suite, a deux points d'inflexion. La courbe de mortalité de Duvillard, dont nous aurons l'occasion de parler plus loin, présente cette particularité que, si l'on réduit les ordonnées à leurs logarithmes, la courbe devient symétrique par rapport à la droite $y = x$; mais, pour cela, il faut choisir l'échelle des nouvelles ordonnées convenablement. Fédor Thoman a construit un certain nombre de courbes de mortalité que l'on peut voir dans un Atlas publié par M. Eugène Pereire.

Vie moyenne, vie probable, etc.

La vie moyenne, à un âge donné, est la moyenne des âges auxquels parviennent les individus de cet âge. Soit $y = f(x)$ le nombre des survivants à l'âge x. Si l'on veut trouver la vie moyenne à l'âge a, il suffira de considérer que dans le temps dx il meurt $-dy$ individus d'âge x : la somme des années qu'ils ont vécu est $-x\,dy$. En faisant ainsi la somme des années qu'ont vécu tous les individus de la Table qui meurent pendant chacun des intervalles tels que dx, on aura la totalité des années pendant lesquelles

ont vécu ces individus, et, en divisant par leur nombre $f(a)$, on aura la vie moyenne à l'âge a ou

$$-\int_a^\infty \frac{x\,dy}{f(a)}.$$

Il vaut mieux écrire cette expression d'une autre manière : en intégrant par parties et en observant que $f(\infty) = 0$, on a

$$-\int_a^\infty \frac{x\,dy}{f(a)} = a + \int_a^\infty \frac{y\,dx}{f(a)}.$$

Cette dernière expression présente sur l'autre l'avantage de ne pas contenir la différentielle de y, laquelle est connue avec moins d'exactitude que y. On conçoit, en effet, qu'une faible variation dans deux ordonnées voisines n'altère pas sensiblement la valeur de y, et modifie cependant beaucoup la direction de la tangente à la courbe de mortalité. Si dans l'expression que nous venons de trouver on fait $a = 0$, on a la vie moyenne à l'époque de la naissance; cette vie moyenne est

$$\int_0^\infty \frac{y\,dx}{f(0)}.$$

Si $f(x)$ représentait la totalité des individus de l'âge x pour une ville ou une nation donnée, $\int_0^\infty y\,dx$ serait la population de cette nation, et la vie moyenne s'obtiendrait, comme l'on voit, en divisant la population par le nombre des naissances. (Dividende et diviseur sont censés constants.)

Laplace, dans sa *Théorie analytique des probabilités*, cherche l'erreur commise dans le calcul de la vie moyenne, en opérant sur un nombre donné de naissances; cette erreur résulte très-simplement de l'erreur commise sur $\frac{f(x)}{f(0)}$, que le théorème de Bayes permet d'évaluer.

La vie probable à un âge donné a est le nombre d'années x après lequel la probabilité que l'individu d'âge a aura d'être vivant sera $\frac{1}{2}$. En d'autres termes, c'est l'âge auquel un individu d'âge a a autant de chances de parvenir ou de ne pas parvenir. Cette vie probable est donc racine de l'équation

$$\frac{f(a+x)}{f(a)} = \frac{1}{2},$$

ou bien

$$2f(a+x) = f(a),$$

facile à résoudre au moyen d'une Table ou d'une courbe de mortalité.

Construction et interpolation des Tables.

Les Tables de mortalité ont surtout pour objet de faire connaître la probabilité de la durée de la vie humaine. Si l'on désigne par $f(x)$ le nombre des survivants à l'âge x, d'après ce que nous avons dit plus haut, $\frac{f(x)}{f(0)}$ est la probabilité qu'a un homme d'atteindre l'âge x, et plus généralement $\frac{f(x)}{f(x')}$ sera la probabilité qu'un individu d'âge x' a d'atteindre l'âge x. En effet, en désignant par u cette probabilité, la probabilité $\frac{f(x)}{f(0)}$ qu'a un individu d'atteindre l'âge x est une probabilité composée de la probabilité $\frac{f(x')}{f(0)}$ qu'il atteindra l'âge x', et de la probabilité u qu'ayant atteint cet âge il atteindra l'âge x. On a donc

$$\frac{f(x)}{f(0)} = \frac{f(x')}{f(0)}u, \quad \text{d'où} \quad u = \frac{f(x)}{f(x')}.$$

Ce raisonnement est fondé, comme l'on voit, sur ce que $\frac{f(x)}{f(o)}$ est la probabilité pour un homme d'atteindre l'âge x; mais cela n'est exact qu'autant que $f(o) = \infty$. Pour l'usage que nous ferons des Tables de mortalité, on peut supposer, sans inconvénient, qu'à la place de $f(o), f(1), \ldots, f(x), \ldots$ on ait mis des nombres proportionnels $F(o)$, $F(1), \ldots$, en sorte que $\frac{F(x)}{F(o)}$ représente la probabilité qu'a un homme d'atteindre l'âge x. En se plaçant à ce point de vue, on remédie à certains inconvénients que présentent les Tables de Deparcieux et de Duvillard, usitées en France, où l'on s'est astreint à écrire le nombre des survivants, ce qui ne donne pas une exactitude constante à toute l'étendue de la Table.

Pour construire une bonne Table, il faut, autant que possible, faire en sorte que le nombre proportionnel $F(x)$ conserve une erreur relative constante, et l'on peut y arriver comme il suit : au lieu de chercher sur $f(o)$ individus combien parviennent aux âges 1, 2, 3,..., on cherchera d'abord sur N individus combien arrivent à l'âge 1, puis sur N individus de l'âge 1 combien arrivent à l'âge 2,.... On connaîtra par ce procédé les rapports $\frac{F(1)}{F(o)}$, $\frac{F(2)}{F(1)}$, $\frac{F(3)}{F(2)}, \ldots$; en appelant $p_1, p_2, p_3, \ldots$ ces rapports, on en conclura, en prenant $F(o) = N$,

$$F(o) = N, \quad F(1) = p_1 N, \quad F(2) = p_1 p_2 N, \ldots .$$

Cette manière de procéder pourrait être employée dans les Compagnies d'assurances, et conduirait à de bons résultats pour la classe d'individus qui s'adressent à elles (*).

(*) Un autre motif nous engage encore à suivre la méthode que nous ve-

Je suppose que l'on ait construit une Table comme nous venons de l'indiquer; on peut chercher à l'interpoler, c'est-à-dire que l'on peut chercher une fonction qui, pour des valeurs de x suffisamment rapprochées, reproduise dans l'intervalle de la Table les nombres $f(x)$ exactement, ou à des quantités près inférieures à l'erreur que le théorème de Bayes assigne aux nombres $f(x)$. Lambert et Duvillard ont proposé les formes suivantes :

$$f(x) = \mathrm{A}\left(1 - \frac{x}{a}\right)^2 + m\left(e^{-\frac{x}{h}} - e^{\frac{x}{k}}\right),$$

où A, a, m, h, k désignent des constantes. Gompertz a proposé une formule fort usitée parmi les actuaires anglais; la voici :

$$f(x) = f(0)\mathrm{G}^{q^x - 1},$$

où G et q sont des constantes. Quoique cette formule contienne moins de constantes que celle de Lambert, elle paraît représenter plus fidèlement le phénomène de la mortalité.

Nous ne terminerons pas ce paragraphe sans faire une

nons d'indiquer. Le théorème de Bayes ne donne des indications précises qu'à la condition que les événements observés ne présentent pas une probabilité trop faible. Une bonne condition à remplir est de faire en sorte que ces probabilités soient voisines de $\frac{1}{2}$. Or en observant sur N individus nés le même jour combien il en survit en cinquante ans, on trouve un rapport $\frac{f(60)}{f(0)}$ assez faible, et les indications du théorème de Bayes sont moins précises. Le même motif engage à ne pas prendre non plus des époques trop rapprochées; parfois l'intervalle d'un an sera trop faible, et il sera bon de calculer F(x) au moyen de la formule

$$\mathrm{F}(x) = \frac{\mathrm{F}(x)}{\mathrm{F}(a)}\,\frac{\mathrm{F}(a)}{\mathrm{F}(b)}\,\frac{\mathrm{F}(b)}{\mathrm{F}(0)}\mathrm{N},$$

où a, b sont des âges convenablement échelonnés.

remarque au sujet de la fonction $\frac{df}{dx}$. La quantité $-\frac{df}{f}$ ou $\frac{f(x)-f(x+dx)}{f(x)}$, ou bien encore $1-\frac{f(x+dx)}{f(x)}$, est la probabilité qu'un individu de l'âge x n'atteindra pas l'âge $x+dx$; car $\frac{f(x+dx)}{f(x)}$ est la probabilité qu'il l'atteindra. Ainsi $-\frac{df}{f}$ est la probabilité de mourir dans l'intervalle de temps dx; donc $-\frac{df}{f\,dx}$ est la proportion des morts dans le temps dx. On a donné à cette quantité le nom de *taux de mortalité;* le taux véritable de mortalité est difficile à évaluer, et de ce que l'on connaît assez bien $f(x)$ il ne faudrait pas en conclure que $f'(x)$ peut être déterminé avec la même précision. En effet, deux fonctions peuvent avoir des valeurs sensiblement égales et néanmoins avoir des dérivées très-différentes.

Probabilité de la durée des associations.

Les Tables de mortalité permettent de calculer la durée des associations, c'est-à-dire de l'existence simultanée de plusieurs têtes. Si nous nous rappelons que $\frac{f(x)}{f(x_0)}$ représente la probabilité qu'un individu d'âge x_0 parviendra à l'âge x,

$$\frac{f(x+a)}{f(a)} \quad \text{et} \quad 1-\frac{f(x+a)}{f(a)}$$

représenteront respectivement la probabilité qu'un individu d'âge a vivra encore x années, et la probabilité qu'il mourra avant x années.

Si l'on considère deux individus d'âges a et b, en vertu

du principe de la probabilité composée,

$$\frac{f(a+x)}{f(a)}\,\frac{f(b+x)}{f(b)}$$

sera la probabilité qu'ils vivront tous deux dans x années;

$$\left[1-\frac{f(a+x)}{f(a)}\right]\left[1-\frac{f(b+x)}{f(b)}\right]$$

sera la probabilité qu'ils cesseront d'exister tous deux dans x années;

$$\left[1-\frac{f(a+x)}{f(a)}\right]\frac{f(b+x)}{f(b)} \quad \text{et} \quad \left[1-\frac{f(b+x)}{f(b)}\right]\frac{f(a+x)}{f(a)}$$

représenteront respectivement la probabilité que dans x années le premier sera mort et le second vivant, et la probabilité que le second sera mort et le premier vivant. Si l'on ajoute ces deux expressions, le principe de la probabilité totale nous apprend que leur somme

$$\frac{f(a+x)}{f(a)}+\frac{f(b+x)}{f(b)}-2\,\frac{f(a+x)}{f(a)}\,\frac{f(b+x)}{f(b)}$$

est la probabilité pour que l'un ou l'autre des individus soit mort dans x années; en d'autres termes, l'expression précédente est la probabilité que dans x années il ne restera plus qu'un seul des deux individus en question;

$$\frac{f(a+x)}{f(a)}+\frac{f(b+x)}{f(b)}-\frac{f(a+x)f(b+x)}{f(a)f(b)}$$

est la probabilité que dans x années l'un des deux individus sera en vie où bien qu'ils seront en vie tous deux, etc.

On pourrait chercher de la même façon la probabilité de la coexistence de trois, quatre, etc. têtes; il faut cepen-

dant ajouter que, si les individus considérés n'étaient pas assujettis à la même loi de mortalité, il faudrait modifier légèrement les résultats qui précèdent. Considérons, par exemple, deux époux : si l'on admet que la mortalité chez l'homme soit $f(x)$ et chez la femme $\varphi(x)$, la probabilité de la durée x d'un mariage entre deux époux d'âges a et b sera

$$\frac{f(a+x)}{f(a)} \frac{\varphi(b+x)}{\varphi(b)}.$$

Dans certaines questions relatives aux assurances sur la vie, on est conduit à remplacer un groupe d'individus par un seul, qui a la même probabilité de s'éteindre à une époque donnée que le groupe en question; on obtient ainsi, dans bien des cas, une approximation suffisante. Mais étudions cette question de plus près et proposons-nous de trouver une loi de mortalité telle que, étant donné un groupe d'individus, il existe toujours un individu possédant la même chance de s'éteindre que le groupe tout entier, quel que soit l'instant où l'on considère le groupe et l'individu isolé.

Soient $a, b, c, \ldots$ les âges des individus du groupe; la probabilité que ce groupe existera dans x années est

$$\frac{f(a+x)}{f(a)} \frac{f(b+x)}{f(b)} \frac{f(c+x)}{f(c)}, \cdots.$$

En égalant ce produit à la probabilité que l'individu d'âge m existera à la même époque ou $\frac{f(x+m)}{f(m)}$, on aura l'équation qui sert à définir $f(x)$ et à calculer m

$$(1) \qquad \frac{f(a+x)}{f(a)} \frac{f(b+x)}{f(b)} \frac{f(c+x)}{f(c)} \cdots = \frac{f(m+x)}{f(m)}.$$

Cette formule doit subsister par hypothèse, quel que soit x;

donc, en prenant la différentielle logarithmique par rapport à x, on aura encore, quel que soit x,

$$(2) \quad \frac{f'(a+x)}{f(a+x)} + \frac{f'(b+x)}{f(b+x)} + \frac{f'(c+x)}{f(c+x)} \ldots = \frac{f'(m+x)}{f(m+x)},$$

ou bien, en posant en général

$$(3) \qquad \frac{f'(u)}{f(u)} = \theta(u),$$

on aura

$$\theta(a+x) + \theta(b+x) + \ldots = \theta(m+x).$$

Cette formule devant avoir lieu, quel que soit x, et pour un groupe de deux individus, par exemple, on aura

$$(4) \qquad \theta(a+x) + \theta(b+x) = \theta(m+x),$$

et, en prenant $x = -m$,

$$\theta(a-m) + \theta(b-m) = \theta(0).$$

Si l'on différentie la formule (4) par rapport à x, et si l'on fait $x = -m$ dans le résultat, on a

$$\theta'(a-m) + \theta'(b-m) = \theta'(0).$$

Posant maintenant $a - m = u$, $b - m = v$, on devra avoir en même temps

$$(5) \qquad \begin{cases} \theta(u) + \theta(v) = \theta(0), \\ \theta'(u) + \theta'(v) = \theta'(0). \end{cases}$$

Différentions alors ces deux équations en considérant v

comme fonction de u, nous aurons

$$\theta'(u) + \theta'(v)\frac{dv}{du} = 0,$$

$$\theta''(u) + \theta''(v)\frac{dv}{du} = 0,$$

et, en éliminant $\frac{dv}{du}$,

$$\frac{\theta''(u)}{\theta'(u)} = \frac{\theta''(v)}{\theta'(v)}.$$

Il est facile de voir que cette relation est une identité : en effet on peut prendre a très-voisin de b; $m-a$ et $m-b$ seront alors très-peu différents (si l'on suppose $f(u)$ continue), et, par suite, u et v seront aussi très-peu différents. On aura donc

$$\frac{\theta''(u+du)}{\theta'(u+du)} = \frac{\theta''(u)}{\theta'(u)} \quad \text{ou} \quad d\frac{\theta''(u)}{\theta'(u)} = 0,$$

c'est-à-dire que $\frac{\theta''(u)}{\theta'(u)}$ est constant; on peut donc poser

$$\theta'(u) = \alpha k e^{-\alpha u},$$

α et k désignant des constantes, et

$$\theta(u) = k e^{-\alpha u} + \text{const.};$$

mais la relation (4) ne peut être satisfaite que si la dernière constante d'intégration est nulle. En effet elle donne

$$k e^{-\alpha(a+x)} + k e^{-\alpha(b+x)} = k e^{-\alpha(m+x)},$$

quand on prend m de telle sorte que l'on ait

$$e^{-\alpha a} + e^{-\alpha b} = e^{-\alpha m},$$

tandis que, si l'on posait

$$e^{-\alpha a} + e^{-\alpha b} + 2\ \text{const.} = e^{-\alpha m} + \text{const.},$$

on n'aurait pas

$$e^{-\alpha(a+x)} + e^{-\alpha(b+x)} + 2\ \text{const.}\, e^{-\alpha x} = e^{-\alpha(m+x)} + \text{const.}\, e^{-\alpha x},$$

si la dernière constante était différente de zéro. Si l'on remplace $\theta(u)$ par la valeur $\frac{f'(u)}{f(u)}$, on en conclut

$$\frac{f'(u)}{f(u)} = ke^{-\alpha u},$$

d'où

$$f(u) = \mathrm{G}\, e^{-\frac{k}{\alpha}e^{-\alpha u}},$$

G désignant une nouvelle constante. Cette formule est au fond celle de Gompertz, et l'on voit ainsi qu'une Table de mortalité interpolée, quand cela sera possible, par la formule de Gompertz, jouira de ce précieux avantage de ramener les opérations sur deux têtes à celles sur une seule tête.

Remarques sur la différentiation et sur l'intégration des fonctions empiriques.

Lorsque l'on a déterminé la forme d'une fonction qui exprime la loi d'un phénomène, on a une tendance généralement trop précipitée à opérer sur cette fonction tout à fait empirique les mêmes calculs que si elle était l'expression rationnelle des faits; en d'autres termes, il semble que l'on oublie que la forme assignée à cette fonction résulte d'hypothèses successives, et que les coefficients n'ont que des valeurs approchées avec une probabilité plus ou moins grande. Le Calcul intégral s'applique en général assez bien

à la fonction empirique qui exprime la loi du phénomène observé; nous dirons tout à l'heure pourquoi, mais le Calcul différentiel peut conduire à des erreurs considérables.

Supposons qu'il s'agisse d'une Table de mortalité et qu'à l'aide de la formule de Gompertz, par exemple, on ait obtenu une fonction interpolatrice telle, que sa différence avec la fonction réelle tombe au-dessous des limites que le théorème de Bayes assigne à l'erreur de la Table. Si l'on déduit le taux de mortalité de la formule de Gompertz en la différentiant, on risque de commettre une erreur à laquelle il semble difficile d'assigner *a priori* une probabilité quelconque : et, en effet, deux fonctions très-peu différentes peuvent avoir des dérivées dont le rapport peut varier dans une proportion considérable. Il suffit, par exemple, que la différence des deux fonctions soit de la forme $\alpha \cos \frac{x}{\alpha^2}$, où α désigne un nombre très-petit; la différence de ces deux fonctions sera moindre que α, tandis que la différence de leurs dérivées $-\frac{1}{\alpha} \sin \frac{x}{\alpha^2}$ pourra devenir égale à $\frac{1}{\alpha}$.

Toutefois, si l'on observe que dans une Table les différences premières, secondes, troisièmes, etc., décroissent rapidement, on pourra interpoler, dans une portion de la Table au moins, à l'aide d'une formule parabolique dont les coefficients auront une probabilité connue, et l'on pourra admettre que l'on calcule avec cette probabilité la dérivée de la fonction qui représente la mortalité au moyen de la formule symbolique

$$\frac{d}{dx} = \Delta - \frac{1}{2}\Delta^2 + \frac{1}{3}\Delta^3 - \ldots,$$

où l'on a supposé, pour simplifier, la différence de la va-

riable x égale à l'unité; cette formule doit être arrêtée dès que les différences deviennent sensiblement nulles. Ce raisonnement ne satisfera certainement pas les esprits rigoureux; mais, dans l'impuissance où nous nous trouvons de découvrir les secrets de la nature, il faut nous contenter de les deviner, quitte à revenir sur nos hypothèses quand elles se trouveront en désaccord avec les faits.

Le Calcul intégral donne des résultats plus précis, et si l'on veut calculer l'aire de la courbe de mortalité, on pourra hardiment faire usage de la formule interpolatrice, ou même des valeurs isolées que fournit la Table de mortalité. En effet, soit y la fonction réelle qui exprime la loi de la mortalité, z la fonction interpolatrice, la différence

$$\int y\,dx - \int z\,dx \quad \text{ou} \quad \int (y-z)\,dx$$

peut être évaluée en substituant à $y - z$ la valeur de la limite de l'erreur ε que le théorème de Bayes assigne à la fonction z. Il y a cependant un cas où l'on peut appliquer la formule

$$\frac{dy}{dx} = \Delta y - \frac{1}{2}\Delta^2 y + \frac{1}{3}\Delta^3 y - \ldots$$

avec sécurité. Je suppose que l'on ait reconnu que dans une grande étendue de la Table $\Delta^3 y = 0$, en considérant chaque valeur de $\Delta^3 y$ comme un événement, la probabilité que $\Delta^3 y$ restera nul dans cet intervalle pour les valeurs intermédiaires de y sera très-grande, et l'on pourra la calculer, par le théorème de Bayes, si Δ^3 est nul; on pourra alors poser avec une probabilité connue

$$\frac{dy}{dx} = \Delta y - \frac{1}{2}\Delta^2 y.$$

Sur l'influence que peut avoir sur la probabilité d'un événement composé l'erreur commise dans l'estimation de la probabilité des événements simples dont il est composé.

Il peut arriver que dans l'estimation que l'on fait de la probabilité d'un événement on commette une erreur, et il est bon de se rendre compte de l'influence que peut exercer cette erreur sur les indications des théories données précédemment. Soit donc p la probabilité approchée d'un événement E et $p \pm \varepsilon$ sa probabilité exacte, la probabilité que E aura lieu α fois dans s épreuves est une fonction de p, que nous connaissons et que nous pourrons désigner par P. Si dans cette fonction on remplace p par $p \pm \varepsilon$, elle devient

$$P' = P + \varepsilon \frac{dP}{dp} + \frac{\varepsilon^2}{1.2} \frac{d^2P}{dp^2} + \ldots$$

ou

$$P'' = P - \varepsilon \frac{dP}{dp} + \frac{\varepsilon^2}{1.2} \frac{d^2P}{dp^2} + \ldots.$$

1° Supposons que l'on connaisse la valeur absolue de ε et que son signe seul soit inconnu, la probabilité que E aura lieu α fois dans s épreuves est un nouvel événement E′, qui est dû à deux causes qui s'excluent mutuellement; ces causes sont les valeurs $p + \varepsilon$ et $p - \varepsilon$ de la probabilité de E. Sa probabilité sera donc, en vertu du principe de la probabilité totale,

$$\frac{1}{2} P' + \frac{1}{2} P'',$$

c'est-à-dire, en vertu des formules précédentes,

$$(1) \qquad P + \frac{\varepsilon^2}{1.2} \frac{d^2P}{dp^2} + \frac{\varepsilon^4}{1.2.3.4} \frac{d^4P}{dp^4} + \ldots.$$

Nous avons supposé tacitement que l'on avait autant de

raisons de croire à la valeur $p-\varepsilon$ qu'à la valeur $p+\varepsilon$. Il est clair que, si ces causes avaient des probabilités différentes, il faudrait modifier un peu notre analyse; en général, c'est la formule (1) qui résoudra la difficulté. On voit que l'influence de l'erreur ε est du second ordre, c'est-à-dire le plus souvent négligeable par rapport à l'erreur ε elle-même.

2° Supposons la valeur de ε inconnue, l'événement E pourra être attribué à chacune des causes qui sont les valeurs de ε comprises entre $-\infty$ et $+\infty$, la probabilité réelle de l'événement E sera donc, en supposant que $\varphi(\varepsilon)\,d\varepsilon$ représente la facilité de l'erreur ε,

$$\int_{-\infty}^{+\infty} P'\varphi(\varepsilon)\,d\varepsilon,$$

ou bien

$$\int_{-\infty}^{+\infty}\left[P\varphi(\varepsilon)+\frac{\varepsilon}{1}\frac{dP}{dp}\varphi(\varepsilon)+\frac{\varepsilon^2}{1.2}\frac{d^2P}{dp^2}\varphi(\varepsilon)+\ldots\right]d\varepsilon.$$

Posant alors, comme plus haut,

$$\int_{-\infty}^{+\infty}\varepsilon^i\varphi(\varepsilon)\,d\varepsilon=M_i,$$

on aura, pour la probabilité cherchée,

$$P+\frac{M_1}{1}\frac{dP}{dp}+\frac{M_2}{1.2}\frac{d^2P}{dp^2}+\ldots.$$

En général, M_1, M_3, M_5 sont nuls et l'expression précédente se réduit seulement à

$$P+\frac{M_2}{2}\frac{d^2P}{dp^2}+\ldots.$$

CHAPITRE V.

SUR LES COMPAGNIES D'ASSURANCES.

Définitions de quelques termes usuels.

Les Compagnies d'assurances ont pour but, comme l'on sait, d'indemniser les particuliers qui contractent avec elles de certains dommages spécifiés à l'avance, moyennant une rétribution en rapport avec les risques connus de part et d'autre.

On distingue plusieurs espèces de Compagnies d'assurances; mais, au point de vue auquel nous nous plaçons dans ce Traité, nous n'aurons à en distinguer que trois espèces : 1° les *Compagnies d'assurances sur la vie*, dont le but est de payer des rentes ou des capitaux dans des circonstances qui dépendent de l'existence d'une ou de plusieurs personnes; 2° les *Compagnies d'assurances sur les choses*, qui garantissent dans des circonstances spéciales le remboursement des pertes occasionnées par le feu, la grêle, les naufrages, etc.; 3° enfin les *Compagnies d'assurances mutuelles et les tontines*, dans lesquelles plusieurs personnes s'associent pour s'entr'aider mutuellement, en payant à frais communs les dommages causés aux associés atteints par certains fléaux spécifiés à l'avance.

Les Compagnies d'assurances sur la vie nous occuperont surtout; nous n'aurons que peu de chose à dire des autres

Compagnies : ce que nous en dirons, du reste, s'appliquera aux Compagnies d'assurances sur la vie.

On donne le nom de *prime* à la somme versée par un assuré pour obtenir l'avantage qu'il attend de la Compagnie. La prime peut être *unique* ou *annuelle, semestrielle, trimestrielle;* la prime unique que l'on paye pour obtenir le payement d'une annuité ou *rente* est ce que l'on appelle le *capital constitutif* de la rente.

Donnons encore le sens de deux mots qui pourront revenir dans le courant de ce Chapitre : l'*actuaire* d'une Compagnie est l'individu chargé de faire les calculs de la Compagnie; le contrat entre l'assuré et la Compagnie porte le nom de *police*.

Considérations sur le taux.

Dans tout ce qui va suivre, nous ferons usage d'une notation uniforme, en nous rapprochant autant que possible de celle qui est en usage parmi les actuaires français : nous désignerons par i le taux de l'intérêt de l'argent.

Mais il y a lieu d'entrer dans quelques détails au sujet de ce que nous appelons le *taux*. Un grand nombre de personnes du monde s'imaginent qu'elles construisent un raisonnement en résolvant une règle d'intérêt simple, et l'on voit malheureusement des professeurs de l'Université s'efforcer de *démontrer* la *règle d'intérêt* en l'assimilant à une règle de trois. De semblables démonstrations sont tout à fait de nature à fausser l'esprit des élèves, et ce que je dis de la règle d'intérêt pourrait se dire également d'un bon nombre d'autres questions. Si l'on désigne par C un capital, c'est-à-dire une somme d'argent, la personne à qui l'on prête cette somme la fait fructifier, et elle s'engage ordinai-

rement à reconnaître le service rendu par le prêteur en lui payant à des intervalles de temps réguliers une somme déterminée, à laquelle on donne le nom d'*intérêt;* on *convient* de déterminer cet intérêt ΔC par la formule

$$(1) \qquad \Delta C = Ci,$$

i désignant un nombre fixé à l'avance et compris ordinairement entre $\frac{3}{100}$ et $\frac{6}{100}$. Rien ne justifie la formule précédente, si ce n'est l'usage. Que l'on ne vienne pas dire que ΔC doit être proportionnel à C, sous prétexte que chaque fraction de capital porte le même intérêt, parce qu'il serait facile de répondre que l'emprunteur ayant besoin d'argent, et le faisant d'autant mieux fructifier qu'il en a davantage, doit reconnaître le service procuré par un capital double en augmentant l'intérêt, de manière à engager le prêteur à lui livrer toute sa fortune, plutôt qu'à la morceler dans diverses affaires. Ainsi la formule $\Delta C = Ci$ ne peut pas et ne doit pas être démontrée : elle est simplement l'expression d'une convention arbitraire. Il en est de même de la formule suivante, où t désigne le temps pendant lequel on laisse le capital à l'emprunteur :

$$(2) \qquad \Delta C = Cit,$$

mais que l'on n'applique guère que lorsque t a une valeur relativement faible, un an ou une fraction d'année.

Il en est de même de la formule

$$(3) \qquad C + \Delta C = C(1 + i)^t,$$

que l'on applique au cas où t représente un nombre entier d'années. Cette dernière formule est celle que nous adopterons pour toutes les valeurs de t, entières ou non, et cela pour plusieurs raisons :

1° Cette formule s'écarte peu de la formule (2) lorsque t est petit; en effet, la formule (3) peut s'écrire

$$C + \Delta C = C\left[1 + it + \frac{t(t-1)}{1.2} i^2 + \ldots\right],$$

ou, aux termes du second ordre près,

$$\Delta C = Cit.$$

2° Elle s'écarte peu aussi, pour la même raison, de la formule enseignée dans les lycées

$$(4) \qquad C + \Delta C = C(1+i)^n(1+fi),$$

où n est le plus grand entier contenu dans t et f la fraction qui, ajoutée à n, donne t.

3° Son calcul est plus facile par logarithmes que celui des autres formules en usage.

4° Elle est employée par les grands financiers et par les actuaires les plus distingués, tels que M. Charlon (*voir* son *Traité des opérations financières*).

5° Le second membre de cette formule, considéré comme l'ordonnée correspondant à l'abscisse t, est représenté dans ses variations par une courbe continue, tandis que les formules (2) et (4) représentent les cordes de cette courbe.

6° Enfin elle se trouve justifiée, sinon démontrée, par le raisonnement qui suit : Si l'emprunteur est toujours assuré de trouver les capitaux qui lui sont nécessaires, il payera pour un temps infiniment petit dt un intérêt $C\tau dt$ proportionnel au capital prêté et au temps dt : on aura donc

$$dC = C\tau dt,$$

d'où l'on tire

$$C = C_0 e^{\tau t},$$

ou, en posant $\tau = 1 + i$,

$$C = C_0 (1 + i)^t,$$

C_0 désignant le capital prêté à l'origine du temps et C ce qu'il est devenu au temps t; i est ce que nous appellerons le *taux*

Formules fondamentales.

Les Compagnies d'assurances doivent être considérées comme jouant avec le public un jeu à peu près équitable. Je dis à peu près, parce qu'elles doivent nécessairement se réserver un léger avantage dans chaque partie jouée, c'est-à-dire dans chacune des affaires qu'elles concluent, afin d'obtenir, dans un grand nombre d'épreuves, un bénéfice certain. Dans toutes les questions relatives aux opérations sur la vie humaine dont nous allons nous occuper, nous égalerons l'importance de la somme espérée par la Compagnie à l'importance de la somme espérée par l'assuré; le bénéfice de la Compagnie se trouvera dans la fixation du taux de l'intérêt de l'argent et dans le choix qu'elle fera de ses Tables de mortalité. Quelques Compagnies bénéficient en augmentant d'une façon arbitraire les résultats calculés; c'est, à notre avis, une façon de procéder tout à fait absurde et sur laquelle nous aurons du reste à revenir un peu plus loin.

Nous allons établir quelques formules fondamentales, en nous servant de notations que nous conserverons dans tout le courant de ce Chapitre : i désignera toujours le taux annuel, en sorte que i soit l'augmentation de 1 franc au bout d'un an; $f(x)$ sera le nombre de survivants à l'âge x, et alors $\frac{f(x)}{f(x')}$ désignera la probabilité qu'une personne d'âge x' a d'arriver à l'âge x (p. 187).

Je résume d'abord dans un petit tableau les formules relatives à la probabilité de la vie humaine, démontrées au Chapitre précédent (p. 187 et 194), et qui nous seront principalement utiles dans la théorie des assurances :

(1) $$\frac{f(a+n)}{f(a)},$$

probabilité qu'une personne d'âge a arrivera à l'âge $a+n$;

(2) $$\frac{f(a+n)}{f(a)}\,\frac{f(b+n)}{f(b)},$$

probabilité que deux personnes d'âges a et b ont de vivre encore ensemble n années;

(3) $$\frac{f(a+n)}{f(a)}\,\frac{f(b+n)}{f(b)}\,\frac{f(c+n)}{f(c)},$$

probabilité que trois personnes d'âges a, b, c ont de vivre encore ensemble n années;

....;

(4) $$1-\frac{f(a+n)}{f(a)} \quad \text{ou} \quad \frac{f(a)-f(a+n)}{f(a)},$$

probabilité qu'une personne d'âge a mourra avant n années;

(5) $$\left(1-\frac{f(a+n)}{f(a)}\right)\left(1-\frac{f(b+n)}{f(b)}\right),$$

probabilité que deux personnes d'âges a et b mourront toutes deux avant n années;

(6) $$\left(1-\frac{f(a+n)}{f(a)}\right)\frac{f(b+n)}{f(b)},$$

probabilité que la personne d'âge b survivra après n années à celle d'âge a;

$$(7)\qquad \frac{f(a+n)}{f(a)}+\frac{f(b+n)}{f(b)}-2\frac{f(a+n)f(b+n)}{f(a)f(b)},$$

probabilité que l'une des deux personnes d'âges a et b sera morte dans n années et l'autre en vie;

$$(8)\qquad \frac{f(a+n)}{f(a)}+\frac{f(b+n)}{f(b)}-\frac{f(a+n)f(b+n)}{f(a)f(b)},$$

probabilité que l'une des deux personnes d'âges a et b au moins sera morte dans n années;

. . . .

Ceci posé, désignons par Q_a^n la valeur actuelle de 1 franc, que l'on ne doit payer que si une personne d'âge a vit au bout de n années, et, par *valeur actuelle*, il va sans dire que nous désignons l'importance de la somme espérée à la mort de la personne en question.

La valeur de Q_a^n n'est autre chose que l'espérance mathématique du franc payable dans n années, si la personne d'âge a est en vie alors. Or, si cette somme n'était soumise à aucune éventualité, elle vaudrait $(1+i)^{-n}$; en la multipliant par la probabilité $\frac{f(a+n)}{f(a)}$ que la personne vivra dans n années, on obtiendra la valeur cherchée: donc

$$(9)\qquad Q_a^n=(1+i)^{-n}\frac{f(a+n)}{f(a)},$$

et il va sans dire que la valeur actuelle d'une somme de c francs payable dans les mêmes conditions sera cQ_a^n.

La valeur actuelle Q_{ab}^n d'une somme de 1 franc payable dans n années, si deux personnes d'âges a et b sont en vie alors, sera l'espérance de cette somme ou le produit de

$(1+i)^{-n}$ par la probabilité (2) : on aura donc

$$(10) \qquad Q_{ab}^{n} = (1+i)^{-n} \frac{f(a+n)}{f(a)} \frac{f(b+n)}{f(b)}.$$

La valeur Q_{abc}^{n} de 1 franc payable dans n années, si trois personnes d'âges a, b, c sont en vie alors, est le produit de $(1+i)^{-n}$ par la probabilité (3), ou

$$(11) \qquad Q_{abc}^{n} = (1+i)^{-n} \frac{f(a+n)f(b+n)f(c+n)}{f(a)f(b)f(c)}.$$

....

La valeur actuelle d'une somme de 1 franc payable dans n années, si une personne d'âge b survit au bout de ces n années à une personne d'âge a, sera $(1+i)^{-n}$ multiplié par la probabilité (6), ou

$$(1+i)^{-n}\left(1 - \frac{f(a+n)}{f(a)}\right)\frac{f(b+n)}{f(b)},$$

ce que l'on peut encore écrire

$$Q_{b}^{n} - Q_{ab}^{n}.$$

La valeur actuelle d'une somme de 1 franc payable dans n années, si l'une quelconque des deux personnes d'âges a, b est morte et si l'autre est en vie, est

$$Q_{a}^{n} + Q_{b}^{n} - 2\,Q_{ab}^{n}.$$

La valeur actuelle d'une somme de 1 franc payable dans n années, si deux personnes d'âges a et b sont mortes toutes deux, ou si l'une des deux est en vie, est

$$Q_{a}^{n} + Q_{b}^{n} - Q_{ab}^{n},$$

....

Des annuités viagères.

On donne le nom d'*annuités viagères* à une série de sommes payables annuellement, tant qu'une ou plusieurs personnes sont en vie :

1° Calculons d'abord la valeur de l'annuité viagère de 1 franc sur une tête d'âge a. Si l'on désigne par C_a cette valeur, elle sera évidemment égale à la somme des espérances mathématiques de 1 franc payable au bout d'un an, deux ans, trois ans, etc., si la personne en question est en vie au bout d'un, deux, trois ans, etc. On a donc, en conservant les notations du paragraphe précédent,

$$(1) \qquad C_a = Q_a^1 + Q_a^2 + Q_a^3 + \ldots,$$

la suite qui entre dans le second membre devant être prolongée jusqu'à ce que l'on rencontre un terme nul provenant de ce que, au bout d'un certain temps, toutes les personnes d'âges a seront certainement mortes. Si l'on désigne par ω l'âge le plus avancé auquel un homme puisse arriver, le dernier terme de la suite sera $Q_a^{\omega-a}$.

Remarque. — Si l'annuité en question était payable d'avance, il faudrait encore ajouter au second membre de la formule (1) le franc payable au commencement de la première année; ainsi l'annuité payable d'avance est $1+C_a$.

2° Calculons maintenant l'annuité viagère sur deux têtes C_{ab}, c'est-à-dire la valeur d'une série de sommes égales à 1 franc, payables tant que deux personnes d'âges a et b seront en vie; la valeur de cette annuité est la somme des espérances mathématiques Q_{ab}^1, Q_{ab}^2, $Q_{ab}^3, \ldots$ de 1 franc, que l'on obtient à la fin de chaque année si les deux personnes en question sont en vie. On a donc

$$(2) \qquad C_{ab} = Q_{ab}^1 + Q_{ab}^2 + Q_{ab}^3 + \ldots,$$

le dernier terme du second membre de cette formule étant de la forme $Q_{ab}^{\omega-a}$ ou $Q_{ab}^{\omega-b}$, selon que a est plus ou moins âgé que b. Nous ferons encore remarquer que, si l'annuité était payable d'avance, sa valeur serait non plus C_{ab}, mais $1+C_{ab}$.

3° Sans qu'il soit nécessaire d'insister beaucoup sur cette question, on voit que l'annuité viagère C_{abc} sur trois têtes est

$$(3) \qquad C_{abc} = Q_{abc}^1 + Q_{abc}^2 + Q_{abc}^3 + \ldots.$$

Les nombres que nous avons désignés par C_a, C_{ab}, C_{abc}, ... sont d'un usage fréquent dans la théorie des assurances ; il importe de montrer comment on peut les calculer. L'application directe des formules (1), (2), (3) serait fastidieuse, et il a fallu songer à les modifier ; plusieurs procédés ont été proposés, et nous allons successivement les faire connaître.

Si l'on veut calculer les valeurs de $C_1, C_2, C_3, \ldots, C_\omega$, pour dresser un tarif d'annuités viagères avec une table de mortalité donnée, on peut procéder comme il suit :

On commence par dresser une Table des nombres que j'appellerai $D_1, D_2, \ldots, D_\omega$, et définis par la relation (*)

$$D_a = f(a)(1+i)^{-a};$$

il est clair que l'on a alors, en vertu de la formule (9) du

(*) Maas, dans son *Traité des annuités viagères*, a préféré diriger le calcul d'une autre manière : il pose

$$S_{a+1} = T_{a+1} + T_{a+2} + T_{a+3} + \ldots, \quad \text{et} \quad T_a = (1+i)^{\omega-a} f(a).$$

On a alors

$$C_a = \frac{S_{a+1}}{T_a}.$$

paragraphe précédent,

$$Q_a^n = \frac{f(a+n)}{f(a)}(1+i)^{-n} = \frac{D_{a+n}}{D_a},$$

et, par suite, la formule (1) devient

$$C_a = \frac{D_{a+1} + D_{a+2} + D_{a+3} \ldots}{D_a},$$

et, en posant

$$E_a = D_a + D_{a+1} + \ldots + D_\omega,$$

il vient

$$(4) \qquad C_a = \frac{E_{a+1}}{D_a}, \quad 1 + C_a = \frac{E_a}{D_a}.$$

Après avoir dressé une Table des nombres D_a, il suffira donc de les ajouter, en commençant par les derniers, pour former une table des nombres E_a, d'où l'on conclura facilement les nombres que nous avons appelés C_a.

Cette méthode peut être remplacée par une autre, qui s'applique aussi avec la même facilité aux annuités sur deux têtes.

Si dans la formule (1) nous changeons a en $a+1$, nous trouvons

$$(A) \qquad C_{a+1} = Q_{a+1}^1 + Q_{a+1}^2 + \ldots;$$

or, si l'on se reporte à la valeur de Q_a^n, on a

$$\begin{aligned} Q_{a+1}^n &= \frac{f(a+n+1)}{f(a+1)}(1+i)^{-n} \\ &= \frac{f(a)}{f(a+1)}(1+i)^{-1} : \frac{f(a+n+1)}{f(a)}(1+i)^{-n-1} \\ &= \frac{f(a)}{f(a+1)}(1+i)\,Q_a^{n+1} = \frac{Q_a^{n+1}}{Q_a^1}. \end{aligned}$$

La formule (A) peut alors s'écrire

$$C_{a+1} = \frac{1}{Q_a^1}(Q_a^2 + Q_a^3 + \ldots),$$

ou bien

$$Q_a^1 C_{a+1} = C_a - Q_a^1,$$

ou bien enfin

(5) $$C_a = Q_a^1(1 + C_{a+1}),$$

formule que l'on aurait pu écrire immédiatement, en observant que l'annuité sur une tête d'âge a est une annuité sur une tête d'âge $a+1$ payable d'avance, ou $1 + C_{a+1}$ ramenée à ce qu'elle serait un an avant le premier versement, à l'aide du multiplicateur Q_a^1. De la formule précédente on tire

$$C_{\omega-1} = Q_{\omega-1}^1(1 + C_\omega), \quad C_{\omega-2} = Q_{\omega-2}^1(1 + C_{\omega-1}), \ldots,$$

et l'on peut faire ainsi le calcul de proche en proche des quantités C_a, quand on a calculé les quantités Q_a^1. Quant à C_ω, il est évidemment nul, en sorte que

$$C_{\omega-1} = Q_{\omega-1}^1.$$

Les quantités C_{ab} se calculent d'une façon analogue; ainsi l'on a

(B) $$C_{a+1,\,b+1} = Q_{a+1,\,b+1}^1 + Q_{a+1,\,b+1}^2 + \ldots.$$

Or on a

$$\begin{aligned} Q_{a+1,\,b+1}^n &= \frac{f(a+n+1)f(b+n+1)}{f(a+1)f(b+1)}(1+i)^{-n} \\ &= \frac{f(a+n+1)f(b+n+1)}{f(a)f(b)}(1+i)^{-(n+1)} \\ &\quad : \frac{f(a+1)f(b+1)}{f(a)f(b)}(1+i)^{-1} = Q_{a,b}^{n+1} : Q_{ab}^1. \end{aligned}$$

La formule (B) devient alors

$$C_{a+1,b+1} = \frac{1}{Q^1_{ab}}(C_{ab} - 1),$$

ou bien

$$C_{ab} = (1 + C_{a+1,b+1})Q^1_{ab},$$

formule analogue à l'équation (5) et qui sert à calculer C_{ab} en fonction de $C_{a+1,b+1}$; on pourra donc dresser une Table de C_{ab}, en observant que $C_{\omega,k} = 0$, quel que soit k, et en calculant une Table des nombres Q^1_{ba} : on aurait de même

$$C_{abc} = (1 + C_{a+1,b+1,c+1})Q^1_{abc},$$

et ainsi de suite.

Les méthodes précédentes sont excellentes pour construire des tarifs, mais elles présentent souvent des inconvénients, auxquels le paragraphe suivant a surtout pour but d'obvier.

Calcul rapide des annuités viagères.

On distingue plusieurs espèces d'annuités : 1° l'annuité payable par année, ou l'annuité proprement dite, que nous avons appris à calculer au paragraphe précédent; 2° l'annuité payable par semestre, par trimestre, par mois, etc.; 3° l'annuité continue.

Une annuité est dite payable par semestre et, en général, par fractions $\frac{1}{n}$ d'année, lorsqu'elle se touche par moitiés ou, en général, par fractions de $\frac{1}{n}$, chaque semestre ou chaque fraction $\frac{1}{n}$ d'année. Ainsi, par exemple, une annuité

de 12 francs, payable par mois, se compose de sommes égales à 1 franc, payables tous les mois.

L'annuité continue est une annuité payable par portions infiniment petites à des intervalles de temps infiniment petits. La considération de l'annuité continue est, comme nous le verrons, de la plus haute importance. Nous désignerons par ${}^{t}C_a$ l'annuité de 1 franc, payable à des intervalles égaux à t et reposant sur une tête d'âge a; ${}^{t}C_{ab}$ sera l'annuité payable à des intervalles égaux à t et reposant sur deux têtes d'âges a et b, etc. En employant cette notation, les annuités continues de 1 franc seront ${}^{0}C_a$, ${}^{0}C_{ab}, \ldots$; les annuités ordinaires seront ${}^{1}C_a$, ${}^{1}C_{ab}, \ldots$.

Calculons ${}^{t}C_a$; en raisonnant comme nous l'avons fait plus haut pour l'annuité ordinaire, il est facile de voir que l'on a

$${}^{t}C_a = t\left[(1+i)^{-t}\frac{f(a+t)}{f(a)} + (1+i)^{-2t}\frac{f(a+2t)}{f(a)} + \ldots\right].$$

Si, dans cette formule, on fait tendre t vers zéro, on trouve

$${}^{0}C_a = \int_0^{\infty}\frac{f(a+x)}{f(a)}(1+i)^{-x}dx;$$

on aurait de même

$${}^{0}C_{ab} = \int_0^{\infty}\frac{f(a+x)f(b+x)}{f(a)f(b)}(1+i)^{-x}dx,$$

et ainsi de suite.

Il est bien difficile de donner des règles précises pour le calcul des quantités C_a, ${}^{0}C_a$, ${}^{t}C_a, \ldots$, dont on a cependant, à chaque instant, besoin. Voici cependant quelques indications générales. Lorsque l'on aura la Table de mortalité

sur laquelle on doit opérer, on commencera par l'*ajuster*, c'est-à-dire par régulariser ses indications, ainsi que nous avons appris à le faire au Chapitre précédent. On observera ensuite que, si l'on construit la courbe, dont l'ordonnée est $\frac{f(a+x)}{f(a)}(1+i)^{-x}$, la quantité ${}^0C_a$ sera l'aire de cette courbe et ${}'C_a$ sera l'aire des rectangles inscrits dans cette courbe, ayant pour base la quantité t. Deux procédés s'offriront alors pour effectuer le calcul de ${}^0C_a$: 1° les méthodes de quadrature, et il conviendra de donner la préférence à celle de Poncelet, qui seule est capable de faire connaître une limite de l'erreur commise dans les calculs; 2° les méthodes d'interpolation. Ces dernières méthodes, discutées avec soin, feront connaître sur quel degré d'exactitude on peut compter, en remplaçant la Table de mortalité par la fonction interpolatrice. Il conviendra, pour faciliter les calculs, de poser

$$f(x) = Ae^{\alpha x} + be^{\beta x} + Ce^{\gamma x} + \ldots.$$

Du reste, la loi de Gompertz, qui est assez bien vérifiée sur certaines Tables, semble indiquer tout naturellement cette forme pour la fonction $f(x)$. En adoptant ainsi la forme exponentielle, l'intégration de $f(a+x)(1+i)^{-x}$ devient très-simple et le calcul de ${}^0C_a$, de ${}^0C_{ab}, \ldots$ devient très-facile pour un taux quelconque; les calculs des annuités ${}'C_a, \ldots$ se réduisent aussi à des sommations de progressions géométriques et il devient inutile de dresser toute une Table d'annuités intermédiaires pour calculer celle dont on a besoin.

Lorsque l'on n'a pas besoin d'une très-grande exactitude, on peut poser, en vertu de la formule d'Euler, qui sert à exprimer les sommes, au moyen des intégrales et *vice versa*;

on a ainsi

$$\int_0^\infty \frac{f(a+x)}{f(a)}(1+i)^{-x}dx = t\sum_{x=0}^{x=\infty}\frac{f(a+x)}{f(a)}(1+i)^{-x} + \frac{t}{2} - \frac{t^2}{12}\frac{f'(a)}{f(a)} + \ldots,$$

ce que l'on peut écrire en se bornant aux premiers termes

$${}^{\circ}C_a = {}^{t}C_a + \frac{t}{2};$$

faisant $t = 1$, on a

$${}^{\circ}C_a = C_a + \frac{1}{2},$$

d'où l'on tire

$${}^{t}C_a = C_a + \frac{1}{2} - \frac{t}{2};$$

telle est la formule approchée dont on fait usage pour calculer les annuités semestrielles et trimestrielles; elle fournit pour l'annuité semestrielle la valeur $C_a + \frac{1}{4}$, pour l'annuité trimestrielle la valeur $C_a + \frac{1}{2} - \frac{1}{8}$ ou $C_a + \frac{3}{8}$; mais l'approximation ainsi obtenue est grossière. On conçoit difficilement la stupidité de quelques assureurs qui, faisant usage de formules aussi peu exactes, se donnent la peine de calculer des annuités avec huit chiffres et les valeurs de leurs contrats en cours avec cet ordre d'approximation ([1])!

([1]) C'est peut-être ici l'occasion de critiquer la manière dont quelques actuaires apprécient les formules qu'on leur soumet. Ces actuaires, qui osent se dire *mathématiciens*, appliquent la formule en question à un exemple numérique; s'ils trouvent le résultat d'accord avec leurs tarifs (qui sont en général faux), ils la regardent comme bonne; dans le cas contraire, ils la

Ce que nous avons dit pour les annuités sur une seule tête peut se répéter à propos des annuités reposant sur un plus grand nombre de têtes; nous ne nous étendrons donc pas plus longtemps sur ce sujet.

(*Voir*, dans le tome I[er] du *Journal des Actuaires*, l'article de M. Achard sur les annuités viagères.)

Formules fondamentales relatives aux assurances sur la vie.

A côté des nombres que nous avons désignés par C_a, C_{ab}, C_{abc}, ..., viennent se ranger une série d'autres nombres, qui jouent un rôle important dans la théorie des assurances et que nous allons apprendre à calculer.

Désignons par P_a la valeur actuelle d'une somme égale à 1 franc que l'on doit payer au décès d'une tête d'âge a, à quelque époque qu'ait lieu le décès. La valeur de P_a est évidemment égale à la somme des espérances mathématiques de sommes égales à 1 franc ramenées à leurs valeurs actuelles et que l'on peut obtenir à une époque quelconque, si le décès de la tête d'âge a a eu lieu à cette époque.

La probabilité que la tête en question s'éteindra au bout du temps x, à une époque comprise entre x et $x+dx$, est le produit de la probabilité $\frac{f(a+x)}{f(a)}$ que cette tête vivra à l'époque x, par la probabilité que, ayant atteint l'âge $a+x$, elle n'atteindra pas l'âge $a+x+dx$ ou

$$1-\frac{f(a+x+dx)}{f(a+x)};$$

---

rejettent sans s'inquiéter de savoir si les méthodes qui y conduisent sont justes et rigoureuses. De telles absurdités se passeraient de commentaires si l'esprit de routine ne les avait profondément ancrées dans les habitudes.

ainsi la probabilité que la tête d'âge a s'éteindra dans le laps de temps considéré dx est

$$\frac{f(a+x)}{f(a)}\left[1-\frac{f(a+x+dx)}{f(a+x)}\right] \quad \text{ou} \quad -\frac{df(a+x)}{f(a)}.$$

En multipliant cette probabilité par la valeur actuelle $(1+i)^{-x}$ de 1 franc, on aura l'espérance mathématique relative à l'événement dont on vient de calculer la probabilité; enfin en ajoutant toutes les espérances analogues relatives aux valeurs de x, comprises entre zéro et la limite extrême de la vie, on aura

$$(1) \qquad P_a = -\int_{x=0}^{x=\infty} \frac{d.f(a+x)}{f(a)}(1+i)^{-x}.$$

Plusieurs méthodes peuvent être employées dans l'évaluation de P_a; la meilleure consiste à intégrer par parties, ce qui donne

$$P_a = \frac{1}{f(a)}\left[f(a) - \int_0^{\infty} f(a+x)(1+i)^{-x}\log(1+i)\,dx\right],$$

ou bien, en développant $\log(1+i)$ par la formule $i - \frac{i^2}{2}\cdots$,

$$P_a = 1 - \int_0^{\infty} \frac{f(a+x)}{f(a)}(1+i)^{-x}dx\left(i - \frac{i^2}{2} + \frac{i^3}{3} - \ldots\right).$$

L'intégrale qui figure dans cette formule est l'annuité continue ${}^0C_a$, en sorte que l'on a

$$(2) \quad P_a = 1 - i\,{}^0C_a + \frac{i^2}{2}\,{}^0C_a - \ldots \quad \text{ou} \quad = 1 - {}^0C_a \log(1+i).$$

Quelques Compagnies appliquent la formule approchée

$$(3) \qquad P_a = 1 - iC_a,$$

et, chose monstrueuse! elles calculent P_a avec quatre chiffres! quelquefois avec huit! L'erreur relative est à peu près $\frac{i^2}{2} C_a$. En prenant $C_a = 10$, terme moyen, et $i = 0{,}04$, on trouve pour cette erreur environ $\frac{1}{100}$.

Il est bon toutefois de se rendre compte de ce que représente le second membre de la formule (3), et de voir dans quelles limites le public ou la Compagnie doit se considérer comme lésé par un contrat d'assurance sur la vie. La quantité P_a est la prime unique que l'on devrait payer à la Compagnie pour obtenir 1 franc le jour du décès de la tête d'âge a; cherchons la prime que l'on devrait payer pour obtenir 1 franc, non plus le jour même du décès de la tête a, mais à la fin de l'année où a lieu le décès.

La probabilité que le décès aura lieu après x années, mais avant $x+1$ années, est, en vertu du principe de la probabilité composée,

$$\frac{f(a+x)}{f(a)}\left[1 - \frac{f(a+x+1)}{f(a+x)}\right] \quad \text{ou} \quad \frac{f(a+x)-f(a+x+1)}{f(a)}.$$

L'espérance mathématique correspondante est

$$\frac{f(a+x)-f(a+x+1)}{f(a)}(1+i)^{-(x+1)},$$

et, par suite, la prime à payer sera

$$(4) \qquad \Pi_a = \sum_{x=0}^{x=\infty} \frac{f(a+x)-f(a+x+1)}{f(a)}(1+i)^{-(x+1)};$$

enfin si (chose absurde en elle-même) la somme devait être payée au commencement de l'année où a lieu le décès, la prime correspondante serait

$$(5)\qquad \Pi'_a = \sum_{x=0}^{x=\infty} \frac{f(a+x) - f(a+x+1)}{f(a)}(1+i)^{-x}.$$

Calculons d'abord Π'_a; on peut écrire

$$\Pi'_a = \sum_0^\infty \frac{f(a+x)}{f(a)}(1+i)^{-x} - (1+i)\sum_0^\infty \frac{f(a+x+1)}{f(a)}(1+i)^{-(x+1)},$$

et, si l'on se reporte à la définition du nombre C_a,

$$\Pi'_a = 1 + C_a - (1+i)C_a,$$

ou enfin

$$\Pi'_a = 1 - iC_a.$$

Ainsi les Compagnies dont nous avons parlé tout à l'heure font payer les primes des assurances comme si les décès avaient lieu au commencement de l'année où ils ont lieu réellement. Les formules (4) et (5) montrent que Π_a est égal à Π'_a au facteur près $1+i$: on a donc

$$\Pi_a = \frac{1 - iC_a}{1+i}.$$

Cette dernière formule pourrait à la rigueur s'appliquer, en prenant pour unité de temps le trimestre, parce que, dans la pratique, il s'écoule toujours un mois ou deux avant que les constatations légales et les formalités à remplir permettent de payer les héritiers d'un assuré; mais alors, en

reprenant l'année pour unité de temps, on devrait prendre

$$P_a = \frac{1 - i^{\frac{1}{4}} C_a}{1 + \frac{1}{4} i};$$

encore cette formule ne serait-elle pas très-exacte, parce que le taux trimestriel n'est pas $\frac{i}{4}$, mais bien $(1+i)^{\frac{1}{4}} - 1$. Du reste, en posant

$$P_a = \frac{1 - n\left[(1+i)^{\frac{1}{n}} - 1\right]^{\frac{1}{n}} C_a}{1 + \frac{1}{n} i},$$

ce qui reviendrait à payer les sommes assurées lorsqu'il s'est écoulé une fraction $\frac{1}{n}$ complète d'année, et en faisant $n = \infty$, on retrouve la formule (2).

Calculons la somme $P_{a,b}$ qu'il faudrait payer actuellement pour obtenir 1 franc au premier décès de deux personnes A, B d'âges a et b.

Si l'on désigne par $d\varphi$ la probabilité que l'une de ces personnes mourra après x années, mais avant $x + dx$ années, l'autre étant encore vivante, on aura, en raisonnant comme tout à l'heure,

$$(6) \qquad P_{a,b} = \int_0^\infty d\varphi(1+i)^{-x}.$$

Reste à calculer $d\varphi$. Cette probabilité est une probabilité totale égale à la somme des probabilités : 1° que A vivra à l'époque x, mais que B mourra à cette époque dans le laps de temps dx; 2° que B vivra à l'époque x, mais que A mourra à cette époque dans le laps de temps dx. La première de

ces probabilités partielles est le produit de la probabilité que A vivra à l'époque $x + dx$ ou $\frac{f(a+x+dx)}{f(a)}$, par la probabilité $-\frac{d.f(b+x)}{f(b)}$ que B mourra dans le laps de temps dx. On trouverait pour la seconde probabilité partielle une expression analogue : on a donc, en négligeant les quantités du second ordre,

$$d\varphi = -\left[\frac{f(a+x)}{f(a)}\frac{d.f(b+x)}{f(b)} + \frac{f(b+x)}{f(b)}\frac{d.f(a+x)}{f(a)}\right],$$

ou bien

$$d\varphi = -\frac{1}{f(a)f(b)}d[f(a+x)f(b+x)].$$

La formule (6) devient ainsi

$$P_{a,b} = -\int_{x=0}^{x=\infty}\frac{1}{f(a)f(b)}d[f(a+x)f(b+x)](1+i)^{-x},$$

ou bien, en intégrant par parties,

$$P_{a,b} = 1 - \int_0^\infty \frac{f(a+x)}{f(a)}\frac{f(b+x)}{f(b)}(1+i)^{-x}\log(1+i)dx,$$

c'est-à-dire

$$P_{ab} = 1 - \log(1+i)C_{ab}.$$

Telle est la véritable valeur de P_{ab}, que l'on remplace parfois par $1 - iC_{ab}$, en commettant ainsi une erreur qui atteint le centième. Les nombres P_{ab}, pas plus que les nombres P_a, ne doivent donc être calculés avec plus de deux décimales quand on fait usage de la formule très-inexacte

$$P_{ab} = 1 - iC_{ab}.$$

On trouverait d'une façon analogue la somme à payer pour obtenir 1 franc au premier décès de 3, 4, ... personnes.

Désignons enfin par $\binom{a}{b}$ la somme qu'il faudrait payer pour obtenir 1 franc au décès de la tête B si la tête A lui survit, les âges respectifs de B et de A étant toujours b et a.

Nous avons vu tout à l'heure que la probabilité que B mourra dans l'intervalle de temps dx et que A vivra encore à cette époque était

$$-\frac{f(a+x)}{f(a)}\frac{d.f(b+x)}{f(b)}=-\frac{f(a+x)}{f(a)}\frac{f'(b+x)}{f(b)}dx.$$

L'espérance mathématique correspondant à cette probabilité s'obtiendra en multipliant par $(1+i)^{-x}$, et l'on aura, par suite,

$$(7)\qquad \binom{a}{b}=-\int_0^\infty \frac{f(a+x)}{f(a)}\frac{f'(b+x)}{f(b)}(1+i)^{-x}dx.$$

Pour évaluer cette intégrale, nous partirons de la formule qui donne l'annuité continue, à savoir :

$$^{\circ}C_{ab}=\int_0^\infty \frac{f(a+x)}{f(a)}\frac{f(b+x)}{f(b)}(1+i)^{-x}dx.$$

On en déduit

$$f(b)\,^{\circ}C_{ab}=\int_0^\infty \frac{f(a+x)}{f(a)}f(b+x)(1+i)^{-x}dx,$$

et, en différentiant par rapport à b,

$$\frac{\partial}{\partial b}[f(b)\,^{\circ}C_{ab}]=\int_0^\infty \frac{f(a+x)}{f(a)}f'(b+x)(1+i)^{-x}dx;$$

la formule (7) devient alors

$$(8)\qquad \binom{a}{b} = -\frac{1}{f(b)}\frac{\partial}{\partial b}[f(b)\,{}^{0}C_{ab}].$$

Telle est la formule exacte qui fournit la valeur de $\binom{a}{b}$; mais on sait que l'on a, en général,

$$\frac{\partial F(b)}{\partial b} = \Delta F(b) - \frac{1}{2}\Delta^2 F(b) + \frac{1}{3}\Delta^3 F(b) - \ldots,$$

la série devant être arrêtée dès que l'on rencontre des différences suffisamment petites. En faisant usage de cette relation, la formule (8) devient

$$\binom{a}{b} = -\frac{1}{f(b)}\left\{\Delta_b[f(b)\,{}^{0}C_{ab}] - \frac{1}{2}\Delta_b^2[f(b)\,{}^{0}C_{ab}] + \ldots\right\},$$

ou, à fort peu de chose près,

$$\binom{a}{b} = \frac{1}{f(b)}[f(b)C_{ab} - f(b+1)C_{a,b+1}]$$

ou

$$\binom{a}{b} = C_{ab} - \frac{f(b+1)}{f(b)}C_{a,b+1}.$$

Si l'on voulait plus d'exactitude, les calculs deviendraient assez pénibles, et ce qu'il y aurait de mieux à faire serait de substituer à l'annuité continue l'annuité mensuelle ou trimestrielle, et à l'année le mois ou le trimestre.

Principe de la composition des contrats.

Toutes les questions relatives aux opérations des Compagnies d'assurances sur la vie peuvent se résoudre à l'aide

d'un principe bien simple, auquel je donnerai le nom de *principe de la composition des contrats*, et qui a beaucoup d'analogie avec le principe de d'Alembert en Mécanique, en ce sens qu'il permet pour ainsi dire d'écrire les résultats d'un calcul sans passer par tous les intermédiaires qu'exigeraient les règles ordinaires du Calcul des probabilités, de même que le principe de d'Alembert permet d'écrire les équations du mouvement sans passer par l'étude individuelle du jeu de chacune des forces du système que l'on soumet à l'analyse.

Le principe de la composition des contrats, que l'on n'apprend bien qu'en l'appliquant fréquemment, a pour but de faire connaître le prix d'une opération financière sur la vie; il peut s'énoncer de la manière suivante :

Pour déterminer le prix d'un contrat, il suffit de le remplacer par d'autres dont les prix sont connus, et qui procurent à l'assuré et à la Compagnie des avantages identiques, et de faire la somme algébrique des prix de tous ces contrats.

Ce principe est aussi évident que le principe de d'Alembert sur le mouvement des systèmes; si l'on ne pense pas à l'appliquer, on peut être entraîné à des calculs fort longs; de même le principe de d'Alembert n'est pas indispensable, mais il simplifie considérablement la mise en équation des problèmes de Dynamique.

Des rentes viagères.

Le mot *rente* est synonyme d'annuité, mais il s'emploie surtout pour désigner l'annuité payée par une Compagnie. La *prime* annuelle, semestrielle, etc., est l'annuité que

reçoit la Compagnie pour un avantage qu'elle procure en échange.

Nous allons montrer comment on peut calculer le prix des différentes espèces de rentes que les Compagnies servent à leurs clients :

1° La rente viagère immédiate est celle que la Compagnie sert à un ou plusieurs individus dès qu'elle a reçu le capital constitutif de cette rente, c'est-à-dire un an après le versement de ce capital si la rente est annuelle, six mois après si elle est semestrielle, trois mois après si elle est trimestrielle, etc. Nous avons appris à calculer le prix de l'annuité viagère sur une ou plusieurs têtes réunies, et nous n'avons rien à ajouter aux détails déjà donnés à cet égard.

2° La rente est *différée de n années* lorsque le bénéficiaire ne la touche que n années après le versement du capital constitutif. Une rente différée d'un an et payable par année est une rente immédiate.

Le prix d'une rente de 1 franc, différée de n années et reposant sur une tête d'âge a, étant représenté par C_a^n, sa valeur dans n années sera celle d'une rente immédiate reposant sur une tête d'âge $a+n$ ou C_{a+n} : sa valeur actuelle est donc $Q_a^n C_{a+n}$, et l'on a (*voir* p. 212)

$$C_a^n = C_{a+n} Q_a^n = C_{a+n} \frac{f(a+n)}{f(a)} (1+i)^{-n} = C_{a+n} \frac{D_{a+n}}{D_a}.$$

Si l'on désigne par C_{ab}^n la valeur de 1 franc de rente, différée de n années et reposant sur deux têtes d'âges a et b, on aura d'une manière analogue (p. 215)

$$C_{ab}^n = C_{a+n,\,b+n} Q_{ab}^n = C_{a+n,\,b+n} \frac{f(a+n) f(b+n)}{f(a) f(b)} (1+i)^{-n},$$

et ainsi de suite.

3° Une rente *temporaire* est une rente dont on ne jouit que pendant un certain temps. Si l'on veut calculer le prix d'une rente temporaire de 1 franc durant n années et reposant sur une tête d'âge a, il suffit d'appliquer le principe de la composition des contrats et de considérer que le bénéficiaire est dans le même état que s'il avait pris une rente viagère C_a en s'engageant à payer une rente différée de n années à la Compagnie dont la valeur actuelle est C_a^n. Le prix d'une rente temporaire de 1 franc reposant sur une tête d'âge a et durant n années est donc

$$C_a - C_a^n.$$

Si la rente reposait sur deux têtes d'âges a et b, elle aurait pour valeur

$$C_{ab} - C_{ab}^n,$$

et ainsi de suite.

4° Une rente de *survie* est une rente sur deux têtes que l'on ne paye qu'au décès d'une des deux têtes. Une rente de survie peut être payable : 1° au survivant désigné; 2° à l'un quelconque des survivants.

Le prix d'une rente de survie égale à 1 franc, reposant sur deux têtes d'âges a et b et payable si la première tête survit à la seconde, s'obtient par le principe de la composition des contrats, en observant que la Compagnie qui sert la rente se trouve dans le même état que si elle avait fait pour ses clients deux contrats, le premier C_a assurant une rente viagère à la tête d'âge a, et le second, $-C_{ab}$, obligeant les assurés à rendre 1 franc de rente tant qu'ils sont en vie tous les deux. Le prix de la rente de survie au profit de la tête d'âge a est donc

$$C_a - C_{ab};$$

la valeur de la rente de survie au profit de l'autre tête est

$$C_b - C_{ab};$$

enfin le prix de la rente payable au premier décès s'obtient en faisant simultanément les deux contrats précédents : son prix est donc

$$C_a + C_b - 2C_{ab}.$$

5° Une rente sur deux têtes A et B, réversible en totalité ou en partie sur l'une des têtes, est une rente que l'on continue à servir en totalité ou en partie à l'une au décès de l'autre.

Pour avoir le prix d'une rente de 1 franc sur deux têtes d'âges a et b, réversible pour la fraction ρ au profit de la première tête, il faut observer que la Compagnie se trouve dans le même état que si elle payait la rente ρ viagère à la tête d'âge a et la rente $1-\rho$ tant que les deux têtes existent ensemble. On a donc pour le prix cherché

$$\rho C_a + (1-\rho)C_{ab}.$$

Cette quantité est aussi égale à une rente de 1 franc pendant l'existence des deux têtes dont le prix est C_{ab}, augmenté d'une rente de survie égale à ρ dont le prix est $\rho(C_a - C_{ab})$. Le prix de la rente au profit de la tête d'âge b serait

$$\rho C_b + (1-\rho)C_{ab}.$$

Enfin le prix de la rente, réversible au profit de l'une ou l'autre tête, s'obtiendra en observant que la Compagnie qui sert la rente se trouve dans le même état que si elle devait payer 1 franc pendant l'existence commune, et une rente de survie de ρ francs au profit de l'un quelconque des survivants ; on doit donc donner à la Compagnie

$$C_{ab} + \rho(C_a + C_b - 2C_{ab}).$$

Si la rente est réversible en totalité, on a $\rho = 1$, et l'expression précédente devient

$$C_a + C_b - C_{ab}.$$

6° Les rentes réversibles peuvent être différées de n années ou être temporaires : le principe de la composition des contrats permettra toujours d'en calculer facilement la valeur; ainsi la rente réversible en totalité et différée de n années a pour valeur $C_a^n + C_b^n - C_{ab}^n \ldots$

Calcul des primes annuelles.

Lorsque l'on a déterminé le prix d'un contrat, en d'autres termes, lorsque l'on connaît la prime unique à payer pour se procurer un certain avantage, on peut se proposer de déterminer la valeur de la prime annuelle qu'il faudrait payer pour obtenir le même avantage. Plusieurs cas peuvent se présenter.

1° *On demande de calculer la prime annuelle x qu'il faudrait payer, pendant toute la vie d'une personne d'âge a, pour remplacer une prime unique V payable immédiatement.*

Cette prime annuelle x n'est autre chose qu'une annuité viagère payable d'avance : sa valeur est donc $x(1 + C_a)$ et, par suite, on a

$$V = x(1 + C_a), \quad \text{ou} \quad x = \frac{V}{1 + C_a}.$$

2° *On demande la prime temporaire qu'il faudrait payer pendant n années, à condition qu'une personne d'âge a soit en vie, pour remplacer une prime unique V payable immédiatement.*

Cette prime temporaire x est une annuité temporaire payable d'avance. Supposons-la d'abord de 1 franc, le prix d'une annuité temporaire de n années payable d'avance est la différence entre le prix d'une annuité viagère payable d'avance, ou $1 + C_a$, et une annuité ordinaire, différée de $n - 1$ années (en vertu du principe de la composition des contrats) : sa valeur est donc

$$1 + C_a - C_a^{n-1};$$

par suite, on a

$$x(1 + C_a - C_a^{n-1}) = V, \quad x = \frac{V}{1 + C_a - C_a^{n-1}}.$$

3° *On demande la prime annuelle équivalant à une prime unique* V *et payable pendant l'existence conjointe de deux têtes d'âge a et b.*

On trouvera, en procédant comme tout à l'heure,

$$\frac{V}{1 + C_{ab}}.$$

4° *On demande la prime temporaire équivalant à une prime unique* V *et payable pendant l'existence simultanée de deux têtes d'âges a et b.*

La prime est payable pendant l'existence conjointe; elle représente une annuité temporaire payable d'avance, mais reposant sur deux têtes, et l'on trouve

$$x = \frac{V}{1 + C_{ab} - C_{ab}^{n-1}}.$$

5° *On demande la prime annuelle équivalant à la prime unique* V *et payable jusqu'au dernier décès de deux personnes d'âges a et b.*

La prime annuelle x est une annuité avec réversion en totalité, mais payable d'avance. Sa valeur pour 1 franc est égale à celle d'une annuité $1 + C_a$ sur la tête a, plus celle d'une annuité $1 + C_b$ sur la tête b, moins celle d'une annuité $1 + C_{ab}$ sur les deux têtes; on en conclut facilement

$$x = \frac{V}{1 + C_a + C_b - C_{ab}}.$$

6° *On demande la prime temporaire de n années équivalant à une prime unique* V *et payable jusqu'au dernier décès de deux personnes d'âges a et b.*

La prime temporaire x est une annuité temporaire payable d'avance; elle est équivalente à une annuité $x(1 + C_a + C_b - C_{ab})$ payable d'avance, tant que vivra l'une des têtes, diminuée d'une annuité différée de $n - 1$ années, payable tant que vivra l'une des têtes ou $x(C_a^{n-1} + C_b^{n-1} - C_{ab}^{n-1})$: on a donc enfin

$$x = \frac{V}{1 + C_a - C_a^{n-1} + C_b - C_b^{n-1} - (C_{ab} - C_{ab}^{n-1})}.$$

Comme on le voit, la prime annuelle capable de remplacer une prime unique pourra toujours se calculer en l'assimilant à une rente payable d'avance et dont le capital constitutif serait la prime unique donnée.

On voit ainsi, par exemple, que la prime temporaire à payer pendant n années, pour obtenir une rente différée de n années, à l'âge a, est

$$\frac{C_a^n}{1 + C_a - C_a^{n-1}};$$

que la prime à payer pendant l'existence conjointe de deux têtes d'âges a et b, pour obtenir une rente de survie

au profit de la tête d'âge a, est

$$\frac{C_a - C_{ab}}{1 + C_{ab}};$$

que la prime à payer pendant n années, pour obtenir une rente de survie au premier décès, la prime cessant d'être exigible alors, est

$$\frac{C_a + C_b - 2C_{ab}}{1 + C_{ab} - C_{ab}^{n-1}},$$

et ainsi de suite.

Des assurances sur la vie.

Les assurances peuvent se partager en deux classes : les assurances en cas de décès et les assurances en cas de vie; nous nous occuperons d'abord des premières.

1° *L'assurance sur la vie entière* est un contrat par lequel la Compagnie s'engage à payer un capital donné au décès d'une ou de plusieurs personnes.

Si l'assurance repose sur une seule tête d'âge a, la prime unique à payer pour 1 franc assuré est le nombre que nous avons désigné par P_a, à la page 223. Cette assurance peut être contractée à primes annuelles et, d'après ce que nous avons vu plus haut, la valeur de chacune d'elles est égale à

$$\frac{P_a}{1 + C_a}.$$

Elle peut aussi se contracter au moyen de primes temporaires payables pendant n années; la valeur de chacune d'elles est alors

$$\frac{P_a}{1 + C_a - C_a^{n-1}}.$$

Si l'assurance repose sur deux têtes et si elle est due au premier décès, la prime unique de cette assurance est la somme que nous avons désignée par P_{ab}; quant à sa prime annuelle payable jusqu'au premier décès, elle est égale à

$$\frac{P_{ab}}{1+C_{ab}}.$$

On pourrait faire les conditions les plus diverses sur le payement de la prime; ce que nous avons dit au paragraphe précédent suffit pour trancher toutes les difficultés qui pourraient se présenter.

Si l'assurance était payable au dernier décès, on trouverait sa prime unique V en observant que, en vertu du principe de la composition des contrats, la Compagnie se trouve dans le même cas, en s'obligeant à payer 1 franc au premier décès et 1 franc au dernier décès, ou en s'obligeant à payer 1 franc au décès de la tête d'âge a et 1 franc au décès de la tête d'âge b. On a donc

$$P_{ab}+V=P_a+P_b,$$

d'où l'on tire

$$V=P_a+P_b-P_{ab}.$$

Connaissant la prime unique V, on en déduit aisément les primes annuelles ou temporaires payables aux diverses conditions qui pourront se présenter.

2° L'*assurance de survie* est un contrat en vertu duquel la Compagnie s'engage à payer une somme déterminée, si une personne d'âge a survit à une autre d'âge b, et cela au moment du décès de la personne d'âge b. Nous avons trouvé le moyen de calculer la prime unique de cette assurance, et nous avons désigné par $\left(\frac{a}{b}\right)$ sa valeur. On en déduit facilement la valeur des primes annuelles équivalentes.

3° L'*assurance temporaire* est un contrat par lequel la Compagnie s'engage à payer une somme déterminée, si une ou plusieurs personnes viennent à mourir, dans un intervalle de temps donné.

Supposons que l'assurance repose sur une seule tête d'âge a et que sa durée soit de n années. Si la Compagnie était obligée d'assurer 1 franc au décès de la tête considérée, quelle que soit l'époque de ce décès, elle devrait recevoir la prime unique P_a; mais cette opération peut se faire : 1° en contractant l'assurance pour n années d'abord, ce qui coûte une somme que je désignerai par V et qui est celle que nous cherchons; 2° en assurant, au bout de ces n années, la tête en question pour le reste de ses jours, ce qui coûtera alors une somme P_{a+n}, somme qui a pour valeur actuelle $P_{a+n}\,Q_a^n$: on a donc

$$P_a = V + P_{a+n}\,Q_a^n,$$

ou bien

$$V = P_a - P_{a+n}\,Q_a^n.$$

Supposons maintenant que l'assurance repose sur deux têtes d'âges a et b, et que la Compagnie s'engage à payer 1 franc au premier décès s'il a lieu avant n années. En raisonnant comme tout à l'heure, on trouve, pour valeur de la prime unique de cette opération,

$$P_{ab} - P_{a+n,\,b+n}\,Q_{ab}^n.$$

Si l'assurance de 1 franc n'est payable que dans le cas où une personne d'âge a survit à une personne d'âge b, dans le délai de n années, la valeur de la prime unique de cette assurance est

$$\binom{a}{b} - \binom{a+n}{b+n} Q_{ab}^n;$$

enfin l'assurance de 1 franc peut être due seulement si les deux personnes meurent dans le délai des n années; la valeur de la prime unique est alors

$$P_a + P_b - P_{ab} - (P_{a+n} + P_{b+n} - P_{a+n,\,b+n})\,Q_{ab}^n.$$

4° Une assurance peut être à *effet différé* et, dans ce cas, la Compagnie n'est engagée que si le décès a lieu après un certain nombre d'années, qui est le *temps du différé*.

Soit n le temps du différé, au bout du temps n, le contrat devient une assurance ordinaire et sa valeur est, par exemple, U; sa valeur actuelle est donc UQ_a^n, $Q_{ab}^n, \ldots$, selon qu'elle repose sur une ou sur deux têtes réunies, etc.

Nous allons maintenant nous occuper des assurances en cas de vie, c'est-à-dire de celles dont l'assuré peut bénéficier lui-même.

1° *L'assurance à terme fixe* est un contrat par lequel la Compagnie s'engage à payer un capital donné à une époque donnée.

Il n'y a pas lieu ici de calculer la prime unique de cette assurance, la somme payée par la Compagnie n'étant soumise à aucune chance aléatoire; si nous prenons, comme d'ordinaire, l'assurance égale à 1 franc et le temps de l'assurance égal à n, la prime unique de cette assurance sera $(1+i)^{-n}$ et sa prime annuelle

$$\frac{(1+i)^{-n}}{1 + C_a - C_a^{n-1}},\quad \frac{(1+i)^{-n}}{1 + C_{ab}^n - C_{ab}^{n-1}},\ldots,$$

selon la manière dont elle devra être payée, tant que vivra une tête d'âge a, tant que vivront deux têtes d'âges a et b, etc.

Dans l'*assurance mixte*, la Compagnie s'engage à payer un capital déterminé à l'assuré, à une époque déterminée, ou à ses héritiers immédiatement en cas de décès.

La valeur de la prime unique de cette opération est la somme de la prime d'assurance temporaire et de l'espérance d'obtenir le capital à la fin du délai. Si a est l'âge de l'assuré, n le délai, la prime d'assurance mixte sera

$$P_a - P_{a+n}\frac{f(a+n)}{f(a)}(1+i)^{-n} + \frac{f(a+n)}{f(a)}(1+i)^{-n};$$

d'où l'on conclut la prime annuelle, en divisant par $1 + C_a - C_a^{n-1}$.

L'assurance de capitaux différés est une opération qui a pour but de se réserver un capital à une époque déterminée. Cette assurance est bonne à faire sur la tête d'une très-jeune fille pour lui procurer une dot à sa majorité; elle s'acquiert alors à l'aide d'une série de primes annuelles faciles à calculer. En effet, en appelant a l'âge de l'enfant, 1 le capital assuré et n le temps du différé, la valeur actuelle du capital est $(1+i)^{-n}\frac{f(a+n)}{f(a)}$; la prime annuelle est une rente payable d'avance et temporaire, dont la valeur pour 1 franc est $1 + C_a - C_a^{n-1}$: la prime cherchée est donc

$$\frac{(1+i)^{-n}\frac{f(a+n)}{f(a)}}{1 + C_a - C_a^{n-1}}.$$

L'assurance temporaire d'annuités est un contrat en vertu duquel la Compagnie s'engage, en cas de décès du contractant, à se substituer à lui pour effectuer un payement d'annuités. Cette assurance est applicable aux personnes qui ont contracté des dettes et qui veulent faire honneur à leurs engagements.

Si l'on désigne par n le nombre des annuités à payer, à l'âge du contractant, la valeur du risque de la Compagnie

est, pour 1 franc d'annuité,

$$(1+i)^{-1}+(1+i)^{-2}+\ldots+(1+i)^{-n}$$

si l'assuré meurt dans la première année, ce dont la probabilité est $\dfrac{f(a+1)-f(a+1)}{f(a)}$,

$$(1+i)^{-2}+\ldots+(1+i)^{-n}$$

s'il meurt dans la seconde, ce dont la probabilité est $\dfrac{f(a+1)-f(a+2)}{f(a)}, \ldots$.

Le risque de la Compagnie, où la prime unique à payer, est donc

$$\frac{f(a)-f(a+1)}{f(a)}[(1+i)^{-1}+\ldots+(1+i)^{-n}]$$
$$+\frac{f(a+1)-f(a+2)}{f(a)}[(1+i)^{-2}+\ldots+(1+i)^{-n}]+\ldots,$$

ou bien, en désignant par A l'annuité temporaire mais certaine de 1 franc,

$$\mathrm{A}-\frac{f(a+1)}{f(a)}(1+i)^{-1}-\frac{f(a+2)}{f(a)}(1+i)^{-2}-\ldots,$$

ou

$$\mathrm{A}-(\mathrm{C}_a-\mathrm{C}_a^n).$$

On retrouve cette valeur en observant que l'assuré paye, en réalité, une rente temporaire sur sa tête et que, par suite, la Compagnie n'a qu'à verser le reste

$$\mathrm{A}-(\mathrm{C}_a-\mathrm{C}_a^n)$$

de l'annuité temporaire certaine; on en conclut, comme dans les cas précédents, la prime annuelle.

La *contre-assurance* est (à mon avis du moins) une fort mauvaise opération, qui a pour but de garantir le remboursement, sans intérêt, d'annuités versées dans une tontine en cas de décès d'une tête sur laquelle les annuités ont été versées. Trois cas peuvent se présenter : 1° les annuités ont été versées; 2° les annuités sont à verser; 3° les annuités sont versées et à verser.

Si les annuités sont versées, la contre-assurance est une assurance temporaire ordinaire et il n'y a rien à ajouter à ce que nous avons dit à ce sujet. On voit que cette opération consiste, en définitive, à jouer deux parties, dont l'une sera nécessairement perdue, ce qui est évidemment plus désavantageux que d'en jouer une seule.

Si les annuités sont à payer, la question prend un aspect différent. Désignons par a l'âge de la tête sur laquelle repose l'assurance, n le temps pendant lequel on verse l'annuité de 1 franc.

L'assurance du premier franc versé est une assurance temporaire de 1 franc pour n années, ou θ_a^n. L'assurance du second franc est encore une assurance temporaire, pour une tête d'âge $a+1$ et pour un temps $n-1$; mais sa valeur actuelle n'est que $\theta_{a+1}^{n-1}\dfrac{f(a+1)}{f(a)}(1+i)^{-1}$, et ainsi de suite.

La valeur de la contre-assurance est donc

$$\theta_a^n + \theta_{a+1}^{n-1}\frac{f(a+1)}{f(a)}(1+i)^{-1} + \theta_{a+2}^{n-2}\frac{f(a+2)}{f(a)}(1+i)^{-2} + \ldots;$$

or on a trouvé, pour l'assurance temporaire θ_a^n, la valeur

$$\theta_a^n = P_a - P_{a+n}\frac{f(a+n)}{f(a)}(1+i)^{-n}.$$

En substituant dans la formule précédente, on a le prix de

la contre-assurance, que l'on peut calculer de plusieurs manières; si l'on n'a pas besoin d'une grande exactitude, on pourra faire usage de la formule approchée

$$P_a = 1 - i C_a.$$

Si la contre-assurance a pour but de rembourser des annuités payées et à payer, elle peut être censée faite en deux contrats et ne présente, par suite, aucune difficulté.

Du rachat.

Il arrive souvent qu'un assuré ne peut plus payer ses primes, ou que, ayant contracté à prime unique, il regrette son contrat. Dans le premier cas, on peut réduire la somme assurée en ayant égard aux primes versées; dans le second, on peut lui permettre de racheter son contrat, en tenant compte des risques courus de part et d'autre.

Si l'assurance est contractée à prime unique, la Compagnie n'a plus rien à toucher; la valeur actuelle de l'assurance est celle d'une assurance reposant sur une tête plus âgée que celle qui est indiquée sur la police, et plus âgée d'un nombre d'années égal au temps écoulé entre la date de la souscription et celle du rachat.

Si l'assurance est contractée à primes annuelles, représentons par a l'âge de l'assuré à la date du contrat, par n le nombre d'années écoulées depuis l'origine du contrat jusqu'au moment de la vente, par p la prime annuelle stipulée dans la police et par k le capital assuré. La valeur des primes qu'il reste à verser est $p(1 + C_{a+n})$; c'est la valeur d'une rente viagère égale à p et reposant sur une tête d'âge $a+n$; d'un autre côté, la valeur de l'assurance, considérée en elle-même ou comme une marchandise vénale, est kP_{a+n}.

L'assuré possède donc le titre positif kP_{a+n} et le titre négatif $p(1+C_{a+n})$; par conséquent la Compagnie doit lui racheter son assurance $kP_{n+n} - p(1+C_{n+n})$. Cette somme est souvent inférieure à la totalité des primes versées.

$$\Sigma[kP_{a+n} - p(1+C_{a+n})],$$

le signe Σ étant étendu à toutes les assurances d'une Compagnie, est ce que l'on appelle sa *réserve;* c'est la somme qu'elle doit tenir disponible et qui représente la valeur de ses contrats.

Le rachat ne se fait que rarement sur les bases équitables dont nous venons de parler, la Compagnie se réserve même le droit de refuser le rachat quand bon lui semble. Elle fait ordinairement imprimer au dos des polices les conditions auxquelles le rachat doit se faire.

Lorsque l'on veut simplement réduire la somme assurée au lieu d'opérer le rachat, on le peut; il suffit pour cela d'observer que la valeur actuelle de l'assurance est

$$kP_{a+n} - p(1+C_{a+n}),$$

et que la somme x que cette valeur est capable d'assurer est donnée par la formule

$$\frac{x}{kP_{a+n} - p(1+C_{a+n})} = \frac{1}{P_{a+n}},$$

d'où

$$x = k - \frac{p}{P_{a+n}}(1+C_{a+n}).$$

Le rachat d'un contrat quelconque se fait d'après les mêmes principes.

Observations générales.

Nous sommes loin d'avoir passé en revue toutes les combinaisons qui pourraient se présenter, mais ce que nous avons dit suffit pour permettre de les résoudre presque toutes.

Il nous reste seulement à faire quelques observations de détail sur la pratique des affaires. Ce qu'il y aurait de plus logique serait d'employer pour les calculs une Table de mortalité construite d'après la nature du public qui s'adresse à la Compagnie dans laquelle on opère. Malheureusement les Compagnies classent leurs polices d'après leurs dates, c'est à dire en définitive sans aucun ordre, en sorte qu'il est fort difficile de faire la statistique, et cette manière de procéder a encore le grand inconvénient d'entraîner à des fautes de calculs dans la préparation des inventaires. A défaut d'une Table construite dans les conditions dont nous venons de parler, on en emploie deux, à savoir : celle de Duvillard et celle de Deparcieux. La première présente une probabilité de mourir plus grande que celle qui est relative aux gens qui s'assurent : on la leur applique. Les contractants sont d'ailleurs visités par un médecin et ne sont assurés que dans le cas où leur santé est bonne.

La Table de Deparcieux, qui a été dressée d'après des observations faites sur des moines, présente une mortalité très-faible et on l'applique au calcul des rentes viagères.

Il est bon de dresser à l'avance des Tables des nombres C_a, D_a, E_a, C_{ab} qui reviennent à chaque instant, et même des nombres ΣE_a dont on peut avoir besoin dans le calcul des contre-assurances, ou même d'autres questions telles que celle-ci, par exemple :

On veut assurer 1 franc au décès d'une personne d'âge a, mais en réservant toutefois à cette personne la facilité de toucher l'intérêt des primes qu'elle verse au taux θ moindre que i. On demande de calculer les primes uniques et annuelles relatives à cette opération.

Pour calculer la prime annuelle, désignons-la par y: la Compagnie reçoit $y(1+C_a)$, l'annuité étant payable d'avance; mais elle donne : 1° 1 franc dont la valeur est P_a; 2° l'intérêt des primes versées, c'est-à-dire la valeur d'une rente progressive d'abord de θy, puis de $2\theta y$, de $3\theta y, \ldots$. En désignant cette rente par $\theta C'_a$, on a

$$y(1+C_a) = P_a + \theta C'_a y,$$

d'où

$$y = \frac{P_a}{1 + C_a - C'_a \theta}.$$

Quant à C'_a, sa valeur est facile à calculer quand on a une Table de C_a; on a évidemment

$$C'_a = (1+i)^{-1}\frac{f(a+1)}{f(a)} + 2(1+i)^{-2}\frac{f(a+2)}{f(a)} + 3(1+i)^{-3}\frac{f(a+3)}{f(a)} + \ldots,$$

ou bien

$$\text{(A)} \quad \left\{ \begin{aligned} C'_a = C_a + C_{a+1}\frac{f(a+1)}{f(a)}(1+i)^{-1} \\ + C_{a+2}\frac{f(a+2)}{f(a)}(1+i)^{-2} + \ldots; \end{aligned} \right.$$

on a du reste

$$C'_{a+1} = C_{a+1} + C_{a+2}\frac{f(a+2)}{f(a+1)}(1+i)^{-1} + \ldots,$$

donc

$$C'_a = C_a + C'_{a+1}\frac{f(a)}{f(a+1)}(1+i)^{-1}.$$

On peut ainsi calculer les C'_a de proche en proche.

Voici un autre moyen de calculer C'_a et qui est plus rapide quand on a une Table des nombres ΣE_a. On a

$$C'_a = \sum_{n=0}^{n=\omega-a} C_{a+n}\frac{f(a+n)}{f(a)}(1+i)^{-n},$$

ou bien, remplaçant C_{a+n} par sa valeur en E et D, ainsi que $\frac{f(a+n)}{f(a)}(1+i)^{-n}$,

$$C'_a = \sum_{n=0}^{n=\omega-a}\frac{E_{a+n+1}}{D_a} = \frac{1}{D_a}\sum_{n=a+1}^{n=\omega} E_n.$$

La prime unique est plus facile à calculer; en la désignant par x, la Compagnie paye une somme qui vaut P_a et une série de sommes égales à θx dont la valeur est $\theta C_a x$: on doit donc avoir

$$P_a + \theta C_a x = x, \quad x = \frac{P_a}{1-\theta C_a}.$$

Si l'on prend $\theta = i$, on a à peu près $x = 1$; car P_a est à peu près égal à $1 - iC_a$, ce qui doit être. Rigoureusement, on a

$$x = 1 + \frac{i^2}{2(1-iC_a)}C_a + \ldots.$$

On conçoit, en effet, que l'assuré jouit du bénéfice de la continuité dont ne jouit pas la Compagnie.

Théorie du plein.

On appelle *plein d'une assurance* la somme maxima qu'un assureur peut accepter sur des risques donnés, sans compromettre ses intérêts; nous allons voir que le plein existe, et nous indiquerons la manière de le calculer approximativement.

Soient $O_1, O_2, \ldots, O_s$ les diverses affaires d'une Compagnie, soient $p_{i1}, p_{i2}, p_{i3}, \ldots$ les probabilités que la Compagnie payera sur l'affaire O_i des sinistres dont les prix, ramenés à leurs valeurs actuelles, soient respectivement a_{i1}, a_{i2}, $a_{i3}, \ldots$. Quelques-unes des quantités p ou a pourront être censées nulles, et Σp devra toujours être égal à l'unité. Si l'on forme alors le produit

$$(1) \qquad f = \Sigma p_{1j} e^{a_{1j}\theta\sqrt{-1}} \Sigma p_{2j} e^{a_{2j}\theta\sqrt{-1}} \ldots \Sigma p_{sj} e^{a_{sj}\theta\sqrt{-1}},$$

le terme en $e^{A\theta\sqrt{-1}}$, dans ce produit, aura pour coefficient la probabilité P que la Compagnie payera à ses assurés des indemnités dont la valeur totale atteindra A. Si A était entier, on pourrait poser

$$(2) \qquad P = \frac{1}{2\pi}\int_{-\pi}^{+\pi} e^{-A\theta\sqrt{-1}} f \, d\theta.$$

Or on pourra toujours supposer A entier, si l'on prend l'unité monétaire infiniment petite, ce que nous ferons, sauf à revenir plus tard à l'unité régulière. Pour calculer la valeur de P, nous mettrons f sous la forme exponentielle, et nous aurons d'abord, au lieu de la formule (1) en observant que la somme des probabilités $p_{i1}, p_{i2}, p_{i3}, \ldots$

est égale à l'unité,

$$f=\left(1+\theta\sqrt{-1}\sum p_{1j}\,a_{1j}-\frac{\theta^2}{2}\sum p_{1j}\,a_{1j}^2+\ldots\right)$$
$$\times\left(1+\theta\sqrt{-1}\sum p_{2j}\,a_{2j}-\frac{\theta^2}{2}\sum p_{2j}\,a_{2j}^2+\ldots\right)\ldots.$$

Nous poserons

(3) $$\Sigma p_{ij}\,a_{ij}=x_i,\quad \Sigma p_{ij}\,a_{ij}^2=z_i;$$

x_i désigne, comme l'on voit, l'espérance de l'assuré O_i, mais non sa prime que l'on prend effectivement un peu plus forte. Nous pourrons alors écrire

$$f=\left(1+\theta x_1\sqrt{-1}-\frac{\theta^2}{2}z_1\ldots\right)\left(1+\theta x_2\sqrt{-1}-\frac{\theta^2}{2}z_2\ldots\right)\ldots,$$

et, par suite,

$$\log f=\sum\left[\theta x\sqrt{-1}-\frac{\theta^2}{2}(z-x^2)\right],$$

en négligeant les termes en θ^3; d'où l'on conclut

$$f=e^{\theta\Sigma x\sqrt{-1}-\frac{\theta^2}{2}\Sigma(z-x^2)}.$$

La formule (2) devient alors

$$P=\frac{1}{2\pi}\int_{-\pi}^{+\pi}e^{-\frac{\theta^2}{2}\Sigma(z-x^2)}\cos(A-\Sigma x)\theta\,d\theta,$$

ou, remplaçant les limites $-\pi$ et $+\pi$ par $-\infty$ et $+\infty$, ce qui n'altère pas beaucoup la valeur de P (p. 34),

$$P=\frac{1}{\sqrt{\pi}}\,\frac{1}{\sqrt{2\Sigma(z-x^2)}}\,e^{-\frac{(A-\Sigma x)^2}{2\Sigma(z-x^2)}}.$$

Si l'on intègre P en faisant varier A depuis $\Sigma x - l$ jusqu'à $\Sigma x + l$, on aura la probabilité que la Compagnie payera aux assurés une somme comprise entre $\Sigma x - l$ et $\Sigma x + l$. On trouvera ainsi, pour cette probabilité,

$$\int_{-l}^{+l} \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{2\Sigma(z - x^2)}} e^{-\frac{u^2}{2\Sigma(z-x^2)}} du;$$

en posant

$$\Theta(\gamma) = \frac{2}{\sqrt{\pi}} \int_0^{\gamma} e^{-\gamma^2} d\gamma,$$

cette expression devient, par un changement de variable facile à deviner,

$$\Theta\left(\frac{l}{\sqrt{2\Sigma(z - x^2)}}\right),$$

et il suffit, comme on le sait, de poser

$$\frac{l}{\sqrt{2\Sigma(z - x^2)}} = \nu \quad \text{ou} \quad l = \nu\sqrt{2\Sigma(z - x^2)},$$

ν étant voisin de 3, pour que Θ soit très-voisin de l'unité : il y aura donc, dans ces circonstances, presque la certitude que la Compagnie payera l'indemnité

$$\Sigma x \pm \varepsilon \quad \text{où l'on a} \quad \varepsilon < \nu\sqrt{2\Sigma(z - x^2)}.$$

La plus grande somme que pourra payer la Compagnie est

$$(4) \qquad \Sigma x + \nu\sqrt{2\Sigma(z - x^2)}.$$

Désignons maintenant par $b_{i1}, b_{i2}, \ldots$ les primes que la Compagnie est à même de recevoir sur l'affaire O_i avec les

probabilités q_{i1}, q_{i2}, ... Si nous posons

$$\Sigma q_{ij} b_{ij} = y_i, \quad \Sigma q_{ij} b_{ij}^2 = u_i,$$

on démontrerait, comme plus haut, que le bénéfice minimum que la Compagnie est en droit d'attendre est très-probablement

$$\Sigma y - \nu\sqrt{2\Sigma(u - y^2)},$$

et, par suite, le gain net de la Compagnie a pour valeur actuelle

$$(5) \qquad \Sigma y - \Sigma x + \nu\left[\sqrt{2\Sigma(z - x^2)} - \sqrt{2\Sigma(u - y^2)}\right].$$

Pour qu'une nouvelle affaire Ω ne soit pas désavantageuse, il faudra qu'en désignant par $\xi, \eta, \zeta, \upsilon$ les valeurs de x, y, z, u relatives à Ω, ces quantités soient telles, que l'accroissement de (5) dû à la nouvelle affaire ne soit pas négatif.

$$\Sigma y + \eta - \Sigma x - \xi + \nu\left\{\sqrt{2[\Sigma(z - x^2) + \zeta - \xi^2]} - \sqrt{2[\Sigma(u - y^2) + \upsilon - \eta^2]}\right\}$$

devra donc être plus grand que (5). En considérant alors $\xi, \eta, \zeta, \upsilon$ comme des infiniment petits, on aura

$$(6) \qquad \eta - \xi > \nu\frac{\zeta - \xi^2}{\sqrt{2\Sigma(z - x^2)}} - \nu\frac{\upsilon - \eta^2}{\sqrt{2\Sigma(u - y^2)}}.$$

Telle est la formule qui servira à décider si l'on doit oui ou non accepter l'affaire Ω. Pour montrer comment on peut faire usage de cette formule, supposons qu'il ne s'agisse que d'assurances vie entière à prime unique : le bénéfice de la Compagnie se composera de la valeur actuelle des primes versées, et η sera la prime unique de l'assurance Ω; cette prime n'est pas égale à ξ, espérance de l'assuré; la

Compagnie, pour se réserver un bénéfice certain, prend $\eta = (1+m)\xi$: m est ce que l'on appelle le *chargement*, on le suppose toujours positif. On a alors, au lieu de l'équation (6),

$$(7) \qquad m\xi > \nu \frac{\zeta - \xi}{\sqrt{2\Sigma(z - x^2)}}.$$

Supposons qu'il s'agisse de calculer le plein d'une assurance vie entière contractée à prime unique sur une tête d'âge c, ξ sera le produit d'une certaine somme α par la prime P_c relative à l'assurance de 1 franc. On aura ensuite

$$\zeta = -\alpha^2 \int_0^\infty \frac{f'(c+x)}{f(c)} (1+i)^{-2x} dx = \alpha^2 P'_c,$$

P'_c désignant la prime d'assurance calculée au taux $2i + i^2$ sur une tête d'âge c, et, par suite, la formule (7) deviendra

$$m P_c \alpha > \nu \frac{P'_c \alpha^2 - P_c^2 \alpha^2}{\sqrt{2\Sigma(z - x^2)}};$$

d'où l'on tire la plus grosse somme que l'on puisse assurer sur une tête d'âge c. Il sera donc essentiel, comme l'on voit, de calculer à l'avance $\sqrt{2\Sigma(z - x^2)}$, tous les ans, par exemple, en faisant l'inventaire de la Compagnie (*).

(*) Nous avons supposé que le bénéfice minimum était indépendant du sinistre maximum à payer, ce qui n'a pas lieu, mais cette hypothèse donne une limite inférieure du plein; si l'on voulait plus d'exactitude, il suffirait de supposer que les sinistres a peuvent être négatifs et représentent alors des bénéfices, mais les calculs numériques seraient alors beaucoup plus compliqués.

NOTE I.

SUR L'INTERPRÉTATION GÉOMÉTRIQUE DE LA MÉTHODE DES MOINDRES CARRÉS.

Supposons qu'il s'agisse de déterminer un point tel que la somme des carrés de ses distances à des plans fixes soit minimum. Soient

$$(1)\qquad \begin{cases} a_1 x + b_1 y + c_1 z = p_1, \\ a_2 x + b_2 y + c_2 z = p_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots, \\ a_s x + b_s y + c_s z = p_s \end{cases}$$

les équations de ces plans, la quantité à rendre minima est

$$\sum \frac{(ax + by + cz - p)^2}{a^2 + b^2 + c^2}.$$

En égalant ses dérivées à zéro, on a

$$(2)\qquad \begin{cases} \sum (ax + by + cz - p) \dfrac{a}{a^2 + b^2 + c^2} = 0, \\ \sum (ax + by + cz - p) \dfrac{b}{a^2 + b^2 + c^2} = 0, \\ \sum (ax + by + cz - p) \dfrac{c}{a^2 + b^2 + c^2} = 0, \end{cases}$$

et l'on voit, en développant ces équations, qu'elles sont identiques à celles que l'on aurait obtenues en appliquant la méthode des moindres carrés aux équations (1), après les avoir d'abord divisées respectivement par $\sqrt{a_1^2 + b_1^2 + c_1^2}$, $\sqrt{a_2^2 + b_2^2 + c_2^2}$, ...

Voici maintenant quelques propositions que nous laisserons démontrer au lecteur. Désignons par M le point en question représenté par les équations (2).

1° Le lieu des points tels que la somme des carrés de leurs distances aux plans (1) soit égale à une constante k^2 est en général un ellipsoïde dont le centre est en M ; quand on fait varier k^2, on obtient des ellipsoïdes homothétiques et concentriques.

2° On peut toujours prendre des coordonnées telles que l'on ait à la fois

$$\sum \frac{ab}{a^2+b^2+c^2} = 0, \quad \sum \frac{ac}{a^2+b^2+c^2} = 0, \quad \sum \frac{bc}{a^2+b^2+c^2} = 0,$$

et cela quelle que soit la position de l'origine.

3° Pour que l'ellipsoïde dont nous venons de parler cesse d'exister, et pour que les équations (2) cessent de représenter un point unique M, il faut et il suffit que les plans (1) passent par des droites parallèles entre elles.

4° Le point M est un point tel que, si de ce point on abaisse des perpendiculaires sur les plans (1), le centre de gravité des pieds de ces perpendiculaires sera précisément le point M.

NOTE II.

SUR LE CALCUL NUMÉRIQUE DANS LES APPLICATIONS DE LA MÉTHODE DES MOINDRES CARRÉS.

Les applications numériques de la méthode des moindres carrés sont généralement très-pénibles à cause de la longueur des calculs qu'elles exigent : il convient donc d'indiquer comment on peut abréger le calcul de Σa^2, Σab, Σac,.... Je suppose que l'on ait à sa disposition une Table de carrés ; les carrés a_1^2,..., b_1^2,... seront donnés par la Table, et l'on commencera par calculer Σa^2, Σb^2, Σc^2,.... On a ensuite

$$\Sigma ab = \frac{1}{4}\Sigma(a+b)^2 - \frac{1}{4}\Sigma(a-b)^2,$$

ou bien encore

$$2\Sigma ab = \Sigma(a+b)^2 - \Sigma a^2 - \Sigma b^2.$$

Ces deux formules sont propres à abréger les calculs; mais la seconde est la plus simple, parce qu'elle n'exige qu'une seule opération compliquée, à savoir, l'addition des nombres $(a+b)^2$, la première exigeant en outre le calcul de $\Sigma(a-b)^2$.

NOTE III.

SUR LE PARAGRAPHE DE LA PAGE 173.

Nous avons cherché la probabilité pour que les erreurs $\xi, \eta, \zeta, \ldots$ des inconnues satisfassent à la relation

$$\sqrt{\left(\frac{\xi-\alpha}{l}\right)^2 + \left(\frac{\eta-\beta}{m}\right)^2 + \ldots} < 1.$$

Géométriquement, cela revenait, dans le cas de trois variables, à faire en sorte que le point ξ, η, ζ, qui est la représentation géométrique de l'erreur, fût compris à l'intérieur d'un ellipsoïde. Si nous avons choisi l'ellipsoïde, c'est qu'il convient en général de calculer des limites de l'erreur différentes pour les diverses variables. Nous aurions pu chercher aussi la probabilité pour que

$$\alpha - l < \xi < \alpha + l, \quad \beta - m < \eta < \beta + m, \ldots;$$

mais ce cas présente quelques difficultés pour le calcul numérique de la probabilité. Toutefois le calcul algébrique est assez simple; le restricteur par lequel il convient de multiplier l'intégrale à évaluer est le produit

$$\left(\frac{1}{\pi}\right)^s \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \ldots e^{\alpha\xi\sqrt{-1}} \frac{\sin ul}{u} du . e^{\beta\eta\sqrt{-1}} \frac{\sin vm}{v} dv \ldots.$$

LISTE
DES
PRINCIPAUX OUVRAGES OU MÉMOIRES
PUBLIÉS
SUR LE CALCUL DES PROBABILITÉS.

Achard (M.). — Méthodes pour le calcul des annuités viagères et des assurances (*Journal des Actuaires*, t. I, 1872).

Adan. — Probabilités du tir, etc. In-8°. — Bruxelles, 1866.

Airy. — On the algebrical and numerical theory of errors of observations. — Cambridge, 1861.

— On the weight to be given to the separate results for terrestrial longitudes determined by observations of transits of the moon and fixed stars (*Mém. de la Société de Londres*, 1851).

Ampère. — Considérations sur la théorie mathématique du jeu. In-4°. — Lyon, 1802.

Andræ. — Fehler-Bestimmung bei der Auflösung der Pothenot'schen Aufgabe, etc. (*Astronomische Nachrichten*, t. XLVII, 1858).

— Udvidelse af en af Laplace *Méc. cél.* angivet Methode for Bestemmelsen af en ubekjend Horrelse ved givne umiddelbare jagtagelser (*Oversigt of det Kgl. danske videnskaberns selskabs Forhandlinger*). — Copenhague, 1860.

Argeländer. — Ueber die Anwendung der Methode der kleinsten Quadrate auf einen besonderen Fall (*Astronomische Nachrichten*, t. XXI).

Bayes. — An essay towards solving a problem in the doctrine of chances. On the late rev. Mr Bayes (*Transact. phil.*, 1763).

Baily. — Théorie des annuités viagères et des assurances sur la vie, suivie d'une collection de tables, 1807. Traduit de l'anglais par M. de Courcy, 1836. In-8°.

Barbier (E.). — Note sur le problème de l'aiguille et le jeu du joint couvert (*Journal de Liouville*, t. V, 1860).

Barrois. — Essai sur l'application du Calcul des probabilités aux assurances contre l'incendie. In-8°. — Paris, 1835.

Beauvisage. — Tables de mortalité et leurs applications. In-8°. — Paris, 1867.

Bellavitis. — Osservazioni sulla theoria della probabilità (*Atti del Istituto veneto di Scienze, Lettere ed Arti*). — Venise, 1857.

Benard. — Note sur une question de probabilité (*Journal de l'Éc. Polyt.*, t. XV, 24e cahier, 1855).

Berkhan. — Ueber die Methode der kleinsten Quadrate. In-8°. — Blankenberg, 1843.

Bernoulli (Daniel). — De duratione mediâ matrimoniorum. (*Novi Comm. Petrop.*, t. XII).

— Essai d'une nouvelle analyse de la mortalité causée par la petite vérole (*Hist. acad.*). — Paris, 1730.

— Specimen theoriæ novæ mensurâ sortis (*Comm. Acad. Petrop.*, t. V).

— De usu algorithmi infinitesimalis in arte conjectandi specimen (*Novi Comm. Petrop.*, 1766, t. XII).

— Disquisitiones analyticæ de novo problemate conjecturali (*Novi Comm. Petrop.*, t. XIV).

— Mensura sortis, etc. (*Novi Comm. Petrop.*, t. XIV).

— Dijudicatio maxime probabilis plurium observationum (*Acta Acad. Petrop.*, 1777).

Bernoulli (Jacqnes). — Ars conjectandi. — Basileæ, 1713. Traduit en français par Vastel. In-8°. — Paris, 1801.

Bernoulli (Jean). — Sur les séquences de la loterie de Genève (*Hist. Acad.*). — Berlin, 1769.

Bernoulli (Nicolas). — Specimina artis conjectandi ad quæstiones juris applicatæ. — Basileæ, 1791.

Bertrand (de Genève). — Développements nouveaux sur la partie élémentaire des Mathématiques (t. I : Probabilités des événements futurs).

Bertrand (J.). — *Voir* Gauss.

Bessel. — Bestimmung des mittleren Fehlers einer Beobachtung (*Astr. Nachrichten*, t. XXI).

— Neue Formeln Jacobi's für einen Anwendungsfall der Methode der kleinsten Quadrate (*Id.*, t. XVII).

— Untersuchung über die Wahrscheinlichkeit der Beobachtungsfehler (*Id.*, t. XV).

— Betrachtungen über die Methode der Vervielfältigung der Beobachtungen (*Id.*, t. XI).

— Ueber das Princip der grossen Zahlen (*Id.*, t. XV).

— Ein Hülfsmittel zur Erleichterung der Anwendung der kleinsten Quadrate (*Astr. Nachrichten*, t. XVII).
— Ueber einen Fehler in der Berechnung der französischen Gradmessung (*Id.*, t. XIX).
— Gradmessung in Ostpreussen. In-4°. — Berlin, 1838.

Bicquiley. — Die Rechnung der Wahrscheinlichkeit. In-8°. — Leipsick, 1778.
— Le même en français. — Paris, 1777.

Bienaymé. — Probabilité des erreurs d'après la méthode des moindres carrés (*Liouville*, t. XVII, 1852).
— Sur la probabilité des résultats moyens des observations, etc. (*Sav. Étrangers*, t. V, 1838).
— Sur la différence qui existe entre la méthode de Cauchy et la méthode des moindres carrés (*Comptes rendus*, 1853).
— Considérations à l'appui de la découverte de Laplace sur la loi des probabilités dans la méthode des moindres carrés (*Comptes rendus*, 1853).
— Sur la loi des grands nombres (*Comptes rendus de l'Acad. des Sciences morales*, 1855).
— Théorème sur la probabilité des résultats moyens des observations (Procès-verbaux de la Soc. Philomathique, 1839).
— Effets de l'intérêt composé (*Id.*, 1839).
— Probabilité de la constance des causes conclue des effets observés (*Id.* 1840).
— Loi de la multiplication et de la durée des familles (*Id.* 1845).

Binet. — Recherches sur une question de probabilité (Th. de Poisson), (*Comptes rendus*, 1844).

Biver. — Théorie analytique des moindres carrés (*Liouville*, t. XVIII, 1853).

Boole. — Proposed questions in the theory of prohabilities (*Cambridge and Dublin Math. Journal*, 1852).

Borchardt. — Ueber Interpolation nach der Methode der kleinsten Quadrate (*Journal de Crelle*, t. LVIII).

Borda. — Mémoire sur les élections au scrutin. In-8°. — Leipzig, 1858.

Bordoni. — Sugli esami, ossia sul merito un esaminato. — Pavia.

Börsch. — Ueber die mittleren Fehler der Resultate aus trigonometrischen Messungen (*Archive de Grunert*, 1866).

Bouniakowski. — Traité du Calcul des probabilités (en russe). — Saint-Pétersbourg, 1846.

Bravais. — Analyse mathématique sur les probabilités des erreurs de situation d'un point.

Bravi. — Theoria e pratica del probabile. — Bergamo, 1840.

Brown (Samuel). — Tables de mortalité, 1858; reproduit, en 1872, dans lo *Journal des Actuaires anglais*.

Brunnow. — Astronomie sphérique et pratique, traduit en français par André et Lucas. — Cet Ouvrage, augmenté par les traducteurs, contient quelques applications de la méthode des moindres carrés. 2 vol. in-8°. — Paris, Gauthier-Villars, 1869-1872.

Buffon. — Essais d'arithmétique morale (4ᵉ vol. du supplément à l'*Histoire naturelle*).

Catalan. — Deux problèmes de probabilités (*Liouville*, t. VI).

— Solution d'un problème de probabilités relatif au jeu de rencontre (*Liouville*, t. II).

Cauchy. — Sur le système de valeurs qu'il faut attribuer à divers éléments déterminés par un grand nombre d'observations. In-4°. — Paris, 1814.

— Sur le système de valeurs qu'il faut attribuer à divers éléments déterminés par un grand nombre d'observations, pour que la plus grande de toutes les erreurs soit un minimum (*Journal de l'Éc. Polyt.*, 20ᵉ Cahier).

— Mémoire sur l'Interpolation (*Liouville*, 1837).

— Divers Mémoires et discussions avec M. Bienaymé au sujet de la méthode des moindres carrés (*Comptes rendus*, 1853).

Casorati. — Intorno ad alcuni punti della teoria dei minimi quadrati (*Annales de Tortolini*, 1858).

Charlon. — Étude sur la tarification des transactions viagères. In-8°. — Paris, 1863.

— Théorie mathématique des opérations financières. In-8°. — Paris, 1869.

Clemens. — Ueber die Methode der kleinsten Quadrate. In-4°. — Tilsit, 1840-1848.

Condorcet. — Mémoires divers sur les prohabilités (*Acad. des Sc.*, 1781, 1782, 1783, 1784).

— Essai sur l'application de l'Analyse à la probabilité des décisions rendues à la pluralité des voix. In-4°. — 1804.

— Article *Assurance* dans l'*Encyclopédie*.

— Sur les événements futurs (*Acad. des Sc.*, 1803).

Costa. — Probabilité du tir. In-8°. — Paris, 1825.

Coste. — Question de probabilité applicable aux décisions rendues par les jurés (*Liouville*, t. VII).

Côtes. — Harmonica mensurarum. — Cambridge, 1722.

Courcy. — *Voir* Baily.

Cournot. — Exposition de la théorie des chances et des probabilités. In-8°. — 1832.

— Sur les applications de la théorie des chances à la Statistique judiciaire (*Liouville*, t. III).

Crofton. — Sur la théorie de la probabilité locale (*Soc. Math. de Londres*, t. II).

— On the theory of probability applied to random straight lines (*Soc. Roy. de Londres*, 1866).

D'Alembert. — *Voir* ses opuscules (in-8°, Paris, 1761). On y trouve traitées diverses questions sur la probabilité; on y voit surtout un grand nombre d'objections.

D'Arrest. — Expressions symétriques pour le poids des inconnues (*Astr. Nachrichten*, t. XLVII).

Dedekind. — Bemerkungen zu einer Aufgabe der Wahrscheinlichkeitsrechnung (*Crelle*, t. L).

— Begründung der Methode der kleinsten Quadrate (*Acad. Berol.*, 1831).

— Mathematische Mitheilungen. In-8°. — Zurich.

Degen (C.-F.). — Tabularum ad faciliorem probabilitatis computationem utilem Enneas. — Kiobenhawn, 1824.

Deparcieux (A.). — Essai sur la probabilité de la vie humaine. In-4°. — 1746.

Deparcieux (neveu). — Traité des annuités. In-4°. — 1781.

Didion. — Calcul des probabilités appliqué au tir des projectiles. In-8°. — 1858.

— Expériences sur la justesse du tir (*Journal de l'Éc. Polyt.*, t. XVI, 27e cahier, 1853).

Dienger. — Ueber die Ausgleichung der Beobachtungsfehler (*Archives de Grunert*, 1852).

— Ueber die Bestimmung des Gewichts der nach der Methode der kleinsten Quadrate, etc. (*Id.*, 1852).

— Ausgleichung der Beobachtungsfehler. Mit zahlreichen Anwendungen. In-8°. — Brunswick, 1857.

Donkin. — An essay on the theory of the combination of observations. In-8°. — Oxford, 1844. Trad. dans *Liouville*, 1850.

Dormoy. — Théorie des jeux (*Journal des Actuaires français*, t. I).

Duhays. — Du jeu de loto (*Liouville*, t. VII).

Duvillard. — Analyse de l'influence de la petite vérole, etc.

— Sa table de mortalité (voir l'*Annuaire du Bureau des Longitudes*).

— Recherches sur les emprunts.

Ellis. — On the foundation of the theory of probabilities (*Trans. of the Cambridge Philosophical Society*, t. VIII).

— On the theory of the least squares (*Id.*).

Enke. — Ueber die Begründung der Methode der kleinsten Quadrat-Summe. In-4°. — Berlin, 1833.

— Ueber die Methode der kleinsten Quadrate (*Berliner astronomisches Jahrbuch*, 1834, 1835, 1836).

Euler. — Probabilité dans le jeu de rencontre. (*Hist. Acad.*). — Berlin, 1751.

— Sur l'avantage du banquier au jeu de pharaon (Acad. de Berlin, 1774).

— Sur la mortalité et les rentes viagères. — Sur la probabilité des séquences à la loterie de Gênes (*Hist. Acad. de Berlin*, 1767).

— Sur la loterie (*Acad. de Berlin*, 1751).

— Sur un théorème de Bernoulli (*Acta Acad. Petrop.*, 1777).

— Opuscula analytica, t. II (sur la loterie et sur les assurances).

— Solution d'une question difficile (*Hist. Acad. de Berlin*, 1767).

— Éclaircissements sur un Mémoire de Lagrange (*Nova Acta Acad. Petrop.*, 1785).

— Cautiones necessariæ in determinatione motus planetarum observandæ (*Id.*).

— Recherches générales sur la mortalité et la multiplication du genre humain (Acad. des Sc. de Berlin, 1740).

Faà de Bruno. — Traité élémentaire du calcul des erreurs. In-8°. — 1869.

Farr (William). — English life tables. — 1864.

Fatio. — Tables d'intérêt, suivies de celles de Buffon et de Halley. — 1778.

Fermat. — Sa correspondance avec Pascal (*OEuvres de Pascal*).

Fischer. — Lehrbuch der höheren Geodäsie. In-8°. — Darmstadt, 1860.

— Theorie der Beobachtungsfehler, etc. In-8°. — Darmstadt, 1845.

Fourier. — Sur les sciences d'observation (*Bulletin de Férussac*, t. II).

— Deux Mémoires sur les résultats moyens déduits d'un grand nombre d'observations. In-4°. — 1826 et 1829.

— Rapport sur les tontines. — 1801.

— Règle usuelle pour la précision des résultats moyens des observations (*Soc. Philomathique*, 1824).

— Sur les compagnies d'assurances (*Annales de Phys. et de Chimie*, 1819).

Freeden. — Die Praxis der Methode der kleinsten Quadrate. In-8°. — Brunswick, 1863.

Fuss. — Recherches sur un problème de probabilité (*Acta Acad. Petrop.*, 1779).

Galloway. — Treatise on probabilities. — Edinburg, 1838.

— On the application of the method of the least squares (*Memoirs of the Astr. Society of London*, 1846).

Gauss. — Theoria combinationis observationum minimis erroribus obnoxiæ (présenté à la Soc. Roy. de Gœttingue, février 1821, le 2 février 1823 et le 16 sept. 1826).

— Même sujet (*Astr. Nachrichten*, t. I, p. 80).

— Même sujet (*Zeitschrift für Astronomie und verwandte Wissenschaften*, t. I, p. 185).

— Sur les éléments de l'orbite de Pallas (*Comm. de Gœttingue*, t. I).

— Même sujet (*Theoria motus corporum cœlestium*).
— Sur la détermination chronométrique des longitudes (*Astr. Nachrichten*, t. V).

N. B. Ces Mémoires ont été traduits par M. J. Bertrand; la traduction se vend chez Gauthier-Villars. In-8°.

Gerling. — Die Angleichungs-Rechnungen der praktischen Geometrie oder der Methode der kleinsten Quadrate. — Hambourg, 1843.
— Sujets analogues (*Astr. Nachrichten*, t. IX et XV).

Goos. — Zur Begründung der Methode der kleinsten Quadrate. In-8°. — Kreutznach, 1865.

Graunt. — Natural and political observations upon the bills of mortality. In-8°. — Londres, 1662.

Grolous. — Sur une question de probabilité appliquée à la théorie des nombres (*Journal l'Institut*, n° 1971), 1872.

Guibert. — Solution d'une question relative à la probabilité des jugements, etc. (*Liouville*, t. III).

Hagen. — Grundzüge der Wahrscheinlichkeitsrechnung. In-8°. — Berlin, 1837 et 1867.

Halley. — An estimate, etc., avec sa Table de Brelaw (*Phil. Trans.*, 1693).

Hansen. — Commentatio de gradus præcisionis computatione. In-8°. — 1830.
— Neue Methode bei Anwendung der Methode der kleinsten Quadrate (*Astr. Nachrichten*, t. VIII).
— Ueber die Anwendung der Wahrscheinlichkeitsrechnung auf geodätische, etc. (*Id.*, t. IX).
— Sur un Mémoire de Gauss relatif aux moindres carrés (*Id.*, t. IX).
— Sur un problème de probabilité relatif à la Géométrie (*Id.*, t. IX).
— Applications de la méthode des moindres carrés à la géodésie (*Mém. de la Soc. des Sc. de Saxe*, 1867, 1868, 1869).

Hauber. — Theorie der mittieren Werthe (*Journal de Baumgärtner*, t. VII).

Hauteserve. — Traité élémentaire sur les probabilités. In-12. — Paris, 1834.
— Application de l'Algèbre élémentaire au Calcul des probabilités. — Paris, 1840.

Hélie. — Mémoire sur la probabilité du tir. In-8°. — Imprimerie Impériale, 1854.

Helmert. — Die Ausgleichungsrechnung nach der Methode der kleinsten Quadrate. In-8°. — 1872.
— Applications du Calcul des probabilités à la géodésie (*Journal de Schlömilch*, 1868 et 1869).

Henka. — Inaugural dissertation (méth. des moindres carrés). In-4°. — Leipsick, 1868.

Herschel. — On theory of probabilities (*Journal des Actuaires anglais*, 1869).

— Sur l'estimation de l'adresse dans le tir (*Annales de l'Observatoire*). Bruxelles, 1866.

Hulltmann. — Sur les moindres carrés. In-4°. — Stockholm, 1860.

Huyghens. — De ratiociniis in ludo aleæ. — 1658.

— Ses œuvres complètes, éditées par S' Gravezende.— Amsterdam, 1728.

Huyn. — Théorie des jeux de hasard. In-8°. — 1788.

Ivory. — On the method of the least squares (*Philos. Magasine*, 1825-1826).

Jahn. — Die Wahrscheinlichkeitsrechnung und ihre Anwendung, etc. — Leipzig, 1839.

— Sur la probabilité du tir et la méthode des moindres carrés. In-8°. — Paris, 1873.

Jordan (C.). — De quelques formules de probabilités (sur les causes) (*Comptes rendus*, 1867).

Jouffret. — Étude sur l'effet utile du tir. In-8°. — Paris, 1873.

— Sur la méthode des moindres carrés et ses applications au tir. In-8°. — 1873.

Jullien (le P.). — Sur la probabilité des erreurs, etc. (*Annales de Tortolini*). — Rome, 1858.

Kerseboom. — Proeven van politique Rekenkunde vervat in drie verhandelingen. In-4°. — La Haye, 1740.

— Sa table de mortalité. — La Haye, 1752.

Knapp. — Ermittelung der Sterblichkeit. — Leipzig, 1868.

Kunzeck. — Studien aus der höheren Physik. In 8°. — Vienne, 1866.

Lacroix. — Traité élémentaire du Calcul des probabilités. In-8°. — 1806.

Lagrange. — Sur l'utilité de prendre le milieu entre les résultats de plusieurs observations (*Misc. Taur.*, 1773).

— Recherches sur les suites récurrentes (*Nouv. Mém. de l'Acad. de Berlin*, 1775).

— Sur une question d'annuités (*Acad. de Berlin*, 1798).

Lambert. — Examen d'une espèce de superstition (*Nouv. Mém. de Berlin*, 1771).

Laplace. — Sur les suites récurro-récurrentes et leur usage, etc. (*Mém. par divers savants*, t. VI, 1774).

— Sur la probabilité des causes par les événements (*Id.*).

— Recherches sur l'intégration des équations aux différencea finiea (*Id.*, 1773).
— Mémoire sur les probabilitéa (*Hist. Acad. de Paris*, 1778).
— Mémoire aur les suites (*Id.*, 1782).
— Sur les mariages, les naissances et lea morts à Paris (*Id.*, 1783).
— Sur les intégrales définies et leura applications aux probabilités, etc. (*Mém. de l'Institut*, 1810).
— Du milieu qu'il faut choisir entre les résultats, etc. (*Connaissance des Temps*, 1813).
— Application du Calcul des probabilitéa à la philosophie naturelle (*Id.*, 1820).
— Application du Calcul des probabilités aux opérationa géodésiques (*Id.*, 1820 et 1822).
— Théorie analytique dea probabilités. Gr. in-4°. — 1812, 1814, 1820, 1847.
— Sur la probabilité du tir (*Mémorial de l'Off. d'Artillerie*, n° 1).

Laurent (H.). — Sur le théorème de J. Bernoulli (*Journal des Actuaires français*, t. I).
— Application du Calcul des probabilités à la vérification des répartitions (*Id.*).
— Sur la conatruction des Tables de mortalité (*Id.*).
— Sur les pleins qu'un assureur peut accepter (*Id.*, t. II).

Legendre. — Nouvelle méthode pour la détermination des orbites des comètes. — Paris, datée 1806, publiée en 1805.
— Méthode des moindrea carrés (*Mém. de l'Institut*, 1810, 1811).

Lemoine. — Solution d'un problème sur les probabilités (*Bulletin de la Soc. math. de Paris*, 1873).

Liagre. — Calcul dea probabilités et théorie des erreurs. In-8°. — Bruxelles, 1852.
— Sur la probabilité d'une cause d'erreur régulière, etc. (*Bull. de l'Acad. de Belgique*, 1855).
— Sur la répartition des hanteurs barométriques (*Id.*, 1852).
— Sur la valeur la plus probable d'un côté géodésique commun à deux triangulationa (*Id.*, 1852).

Lindenau. — Sur l'orbite probable de la comète de 1680 (*Journal de Lindenau*, 1818).

Littrow. — Die Wahrscheinlichkeitsrechnnng in ihrer Anwendung auf das wissenschaftliche und praktische Leben. In-8°. — Vienne, 1832.

Lobatchewsky. — Probabilité des résultats moyena d'observations répétées (*Crelle*, 1824).

Lobatto. — Remarques sur nne formule fausse de M. Reboul pour évaluer le prix d'une assurance de snrvie (*Archives néerlandaises*, t. I, 1866).

Loua. — Table de mortalité, dans la Statistique de la France, publiée par le Ministère de l'Agriculture et du Commerce. Gr. in-4°. — 1873.

Maas. — Théorie élémentaire des annuités viagères et des assurances sur la vie.

Macleroth. — Mémoire à consulter sur la marche et la succession des événements aléatoires. — Paris, 1840.

Makeham. — On the integral of Gompertz's fonction... (*Journal des Actuaires anglais*, t. XVII).

Malfatti. — Sur un problème de D. Bernoulli (*Mém. de la Soc. italienne*, t. I, 1782).*

Mallet. — Sur le Calcul des probabilités (*Acta helvetica*, t. VII).

Mertz. — Die Wahrscheinlichkeitsrechnung und ihre Anwendung, etc. In-8°. — Frankfurt, 1854.

Meyer. — Mémoire sur l'application du Calcul des probabilités aux opérations du nivellement topographique en 1848.

— Essai d'une exposition nouvelle de la théorie analytique des probabilités *a posteriori*. In-4°. — Liége, 1857.

Moheau. — Recherches sur la population. — 1778.

Moivre. — *Voir* les *Miscellanea analytica*. — 1730.

— Doctrine of chances. In-4°. — 1756, 1716, 1738.

— De mensurâ sortis (*Phil. Trans.*, 1711).

— Annuities on lives. In-8°. — 1724, 1742, 1750, 1752.

Mondésir. — Solution d'une question qui se présente dans le Calcul des probabilités (*Liouville*, t. II).

Montferrand (de). — Table de mortalité (26e Cahier du *Journal de l'École Polytechnique*).

Montmort. — Analyse des jeux de hasard. — 1713.

Morgan. — The principles and doctrine of assurances. In-8°. — 1821.

— On the theory of observations (*Transact. de la Soc. Phil. de Cambridge*, 1864).

Neovius. — Larobok in minsta Qvadrat metoden. In-8°. — Abo, 1870.

Newton. — Tables d'intérêt. 6e édit. — Londres, 1742.

Nicole. — Méthode pour déterminer le sort des joueurs (*Hist. Acad.*) — Paris, 1730.

— Résolution de quelques questions sur les jeux (*Id.*).

Noble. — The application of theory of probabilities to artillery praktice (*Occasional Papers of Roy. Art. Instit.*, t. I).

Oppolzer. — Expressions symétriques pour le poids des inconnues (*Astr. Nachrichten*, t. LXVII).

Ostrogradsky. — Probabilité des jugements (*Acad. de Saint-Pétersbourg*, 1834).
— Sur la probabilité des hypothèses (*Mélanges math. et astr.*, 1859).

Ottinger. — Die Wahrscheinlichkeitsrechnung. — Berlin, 1852.
— Sur le jeu de rouge et noire (ou trente et quarante) (*Crelle*, t. LXVII).

Otto (colonel). — Erörterungen über die Mittel für Beurtheilung der Wahrscheinlichkeit des Treffens (*Archiv für Offiziere des Kgl. preussischen Artillerie- und Ingenieurkorps*, t. XXXIX et XL).

Parisot. — Traité de calcul conjectural. In-8°. — Paris, 1810.

Pascal. — Diverses questions sur le Calcul des probabilités, lettres de Pascal et Fermat (*voir* ses Œuvres, 3e vol., publiées par Drion, chez Hachette, 1862).

Patavio. — Probabilismus methodo mathematica demonstratus. — 1840.

Peaucellier et **Wagner.** — Sur la théorie des erreurs appliquées à un nouvel appareil autoréducteur (*Mémorial de l'Off. du Génie*, n° 18, 1868).

Paucker. — Ueber die Anwendung der Methode der kleinsten Quadrate und physikalische Beobachtungen. In-4°. — Mittau, 1819.
— Note zur Theorie der kleinsten Quadrate (*Bull. de l'Acad. des Sc. de Saint-Pétersbourg*, 1850-1851).

Pereire (E.). — Tableaux sur les questions d'intérêt et de mortalité. Atlas in-folio. — Paris, 1869.
— Tables d'intérêts composés et d'assurances viagères. — Paris, 1872.

Peters. — Ueber die Bestimmung des wahrscheinlichen Fehlers (*Ast. Nachrichten*, 1822).
— Wahrscheinlicher Fehler einer Beobachtung, etc. (*Id.*).
— Wahrscheinlichste Elemente der Bahn des Kometen von 1585 (*Id.*, 1829).

Plana. — Sur la théorie des erreurs (*Acad. de Turin*, 1811-1812).

Plœnnis (de). — De la distinction à établir entre les diverses données de la dispersion et de leur usage pratique pour la détermination de la probabilité du tir (*Journal des Armes spéciales*, 1864).

Pochet. — Sur le jeu de l'horloge (*Journal des Actuaires français*, t. II).
— Géométrie des jeux de bourse (*Id.*).

Poisson. — Recherches sur la probabilité des jugements, etc., précédées des règles générales du Calcul des probabilités. In-4°. — 1837.
— Solution d'un problème de probabilité (*Liouville*, t. II).
— Sur les naissances (*Acad. des Sc.*, t. IX).
— Sur le jeu de trente et quarante (*Gergonne*, t. XVI).
— Loi des grands nombres (*Comptes rendus*, 1835).
— Sur la proportion des naissances des deux sexes (*Conn. des Temps*, 1827).

— Sur la probabilité du tir (*Mémorial de l'Off. d'Artillerie*, 1829, n° 3; *voir* aussi le n° 2).

— Sur la probabilité des résultats moyens des observations (*Conn. des Temps*, 1827-1832).

Poudra et **Hossard.** — Qnestion de probabilité résolue par la Géométrie, In-8°. Paris, 1819.

Prévost. — Sur les principes de la théorie des gains fortuits (*Nouv. Mém. de Berlin*, 1780).

Price. — Observations on reversionary payments. — 1812.

— Sur le théorème de Bayes (*Transact. Phil.*, 1764 : « A demonstration of second rule in the essay, etc. »).

Puissant. — Article *Équations de condition* (*Dict. des Math.* de Montferrier, Supplément).

— Application du Calcul des probabilités à la mesure de la précision. In-8°. — Paris, 1830.

— Mémoires sur l'application du Calcnl des probabilités aux mesures géodésiques (*Acad. des Sc.*, 1831 et 1832, et *Comptes rendus*, 1838).

— *Voir* le Supplément au *Traité de Géodésie*.

Quetelet. — Théorie des probabilités. In-8°. — Bruxelles, 1853.

— Lettres sur la théorie des probabilités (adressées au duc de Saxe-Cobourg). In-8°. — Bruxelles, 1845.

— Recherches sur la fécondité et la mortalité en Belgique. (En collaboration avec M. Smith). In-8°. — 1858.

— Essai de physique sociale. — Paris, 1834.

— Snr qnelques propriétés curieuses que présentent les résultats d'une série d'observations, etc. (*Acad. de Belgique*, 1852).

— Dictionnaire d'économie politique, par Coquelin et Guillaumin (*voir* l'article Table des mortalités. In-4°. — Paris, 1853.

Ramus. — Sur une question de probabilité relative aux corrections des hauteurs barométriques (*Crelle*, t. XXIV).

Regnault. — Calcul des chances et philosophie de la Bourse. In-8°. — Paris, 1863.

Reuschle. — Sur les moindres carrés (Ueber die Deduction, etc., *Crelle*, t. XXVI et XXVII).

Ritter. — Manuel théorique et pratique de l'application de la méthode des moindres carrés. In-8°. — Paris, 1856.

Roberts. — An arithmetical paradox concerning the chance of loteries (*Phil. Trans.*, 1693).

Roger. — Solution d'un problème de probabilité (*Liouville*, t. XVII).

Rosemberger. — Mittheilung der Bessel'schen Methode, etc. (*Astr. Nachrichten*, t. VI).

Ruffini. — Réflexions critiques sur l'essai philosophique de Laplace (en italien). — Modène, 1821.

Sawitsch. — Anwendung der Wahrscheinlichkeitstheorie. — Leipzig, 1863.

Scheffler. — Sterblichkeit und Versicherungswesen neu behandelt. — Braunschweig, 1868.

Schindler. — Ueber Fehler bei der Berechnung eines ebenen Dreiecks. In-4°. — Prague, 1858.

Simon. — Exposition des principes du Calcul des probabilités (*Journal des Actuaires français*, t. I).

Simpson (Th.). — The nature and laws of chance. In-8°. — 1740.

— The doctrine of annuities and reversions. In-8°. — 1742.

— An attempt to show the advantage arising by taking the mean of a number of observations in practical Astronomy (*Miscellaneous tracts on some curious subjects*, 1757).

Smart. — Table de mortalité pour Londres, 1742.

Strodtmann. — Bevathelijk onderrigt in de Kansrekening. — Breda, 1834.

Süssmilch. — Table de mortalité, corrigée par Baumann, dans l'Ouvrage intitulé : *Die göttliche Ordnung*. In-8°. — 1775.

Tchebycheff. — Démonstration d'une proposition importante, etc. (théorème de Bernoulli), (*Crelle*, t. XXXIII).

— Sur l'interpolation (*Acad. de Saint-Pétersbourg*, 1853-1864).

— Essai d'une analyse élémentaire de la théorie des probabilités. In-8°.

— Des valeurs moyennes (*Liouville*, 1867).

Tedenat. — Sur la loterie (*Gergonne*, t. XII).

Todhunter. — A history of the mathematical theory of the probability. In-8°. — London, 1865 (*).

Tozer. — On the measure of the force of testimony (*Transact. Phil. de la Soc. de Cambridge*, t. VIII).

Trembley. — Sur la durée des mariages (*Acad. de Berlin*, 1794-1795).

— Sur la méthode de prendre les milieux entre les observations (*Id.*, 1801).

— Observations sur le calcul d'un jeu de hasard (*Id.*, 1802).

— Disquisitio circa calculum probabilium (*Comm. de Gœttingue*, t. XII).

— De probabilitate causarum ab effectibus oriunda (*Id.*, t. XIII).

— Recherches sur une question relative aux probabilités (*Acad. de Berlin*, 1794-1795).

(*) Cet Ouvrage nous a été très-utile pour la construction de cette Table bibliographique; il contient un très-grand nombre de documents.

Verdam. — Verhandeling over de Methode der kleinsten Quadrate. In-8°.

Violeine. — Tables pour faciliter les calculs des probabilités sur la vie humaine. In-4°. — Paris, 1859.

Vorlander. — Ausgleichung der Fehler polygonometrischer Messungen. In-8°. — Leipzig, 1858.

Wagner. — *Voir* Peaucellier.

Wargentin. — Table de mortalité (*Mémoires de l'Acad. de Stockholm*, 1776).

Waring. — An essay on the principles of human knowledge. — 1794.

Wild. — Die Grundsätze der Wahrscheinlichkeitsrechnung. In-8°. — Munich, 1862.

Winckler. — Allgemeine Sätze zur Theorie der unregelmässigen Beobachtungsfehler (*Comptes rendus de l'Acad. des Sc. de Vienne*, 1866).

Windelband. — Die Lehren vom Zufall. — Berlin, 1870.

Witt (Jean de). — Annuités viagères (ses Œuvres trad. en français à La Haye, 1709).

Wittstein. — Mathematische Statistik. In-4°. — Berlin, 1867.

Wolff. — Note zur Methode der kleinsten Quadrate (*Mém. de la Soc. des Sc. nat. de Berne*, n°s 160-161).

— Versuche zur Vergleichung der Erfahrungswahrscheinlichkeit, etc. (*Id.*, n°s 156 à 269).

Woolhouse. — On the philosophy of statistics (*Almanach for the year* 1872).

— *Voir* divers articles du *Journal des Actuaires anglais*, sur l'ajustement des Tables et les annuités (t. XII et XV).

— On the philosophy of statistics (*Journal des Actuaires anglais*, t. XVII).

Young. — Remarks on the probabilities of errors in physical observations (*Phil. Trans.*, 1819).

Yvon Villarceau. — Sur la méthode d'interpolation de Cauchy (*Conn. des Temps*, 1852).

Zech. — Eine Abhandlung zur Methode der kleinsten Quadrate.

Zeuner. — Abhandlungen aus der mathematischen Statik. — Leipzig, 1869.

FIN.

532 Paris. — Imprimerie de GAUTHIER-VILLARS, quai des Augustins, 55.


Digitized by Google



